

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 04:02:02 ; Search time 3845 Seconds  
(without alignments)  
10914.498 Million cell updates/sec

Title: US-09-938-703-3  
Perfect score: 1442  
Sequence: 1 GAAATCCCAACAGAGCCA.....AGTAGTAGATTCGGAATTC 1442

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenBank:

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vi.\*
- 15: en.ba.\*
- 16: en.fun.\*
- 17: en.hum.\*
- 18: en.in.\*
- 19: en.ru.\*
- 20: en.om.\*
- 21: en.or.\*
- 22: en.ov.\*
- 23: en.pat.\*
- 24: en.ph.\*
- 25: en.pl.\*
- 26: en.ro.\*
- 27: en.sts.\*
- 28: en.un.\*
- 29: en.vi.\*
- 30: en.htg.hum.\*
- 31: en.htg.inv.\*
- 32: en.htg.other.\*
- 33: en.htg.mus.\*
- 34: en.htg.pln.\*
- 35: en.htg.rod.\*
- 36: en.htg.man.\*
- 37: en.htg.vrt.\*
- 38: en.sy.\*
- 39: en.htgo.hum.\*
- 40: en.htgo.mus.\*
- 41: en.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1442	100.0	1442	6	AX427283 Sequence
2	1417.2	98.3	7422	9	AF009962 Homo sapi
3	1382.8	95.9	1477	6	AX153803 Sequence
4	1382.8	95.9	1477	6	AX427281 Sequence
5	1375.2	95.4	6059	9	HSCCR5AB2
6	1375.2	95.4	143068	6	AX335952 Sequence
7	1375.2	95.4	143068	9	HSU95626
8	1375.2	95.4	185437	9	AC098613
9	1344	93.2	1344	6	AR119948 Sequence
10	1302	90.3	1376	6	AR119946 Sequence
11	1302	90.3	1376	9	HSCCKR4G
12	1293.4	89.7	1414	6	AX297794 Sequence
13	1293.4	89.7	1414	6	AX297803 Sequence
14	1293.4	89.7	1414	6	AX297812 Sequence
15	1293.4	89.7	1414	6	AX297821 Sequence
16	1293.4	89.7	1414	6	AX297830 Sequence
17	1293.4	89.7	1414	6	AX297839 Sequence
18	1288.6	89.4	1414	6	BD006753
19	1283.2	89.0	1966	9	MMU77672
20	1149	79.7	3383	6	AR161694
21	1149	79.7	3383	6	BD006760
22	1149	79.7	3383	9	HSU54994
23	1117	77.5	1225	6	AX256200
24	1117	77.5	1225	9	HSU57840
25	1115.4	77.4	1225	6	AX256202
26	1027	71.2	1027	9	HSCMKBR5
27	1027	71.2	1027	9	HSU66285
28	1012	70.2	1086	9	HSU83326
29	1005.2	69.7	1092	9	U97666
30	997.2	68.2	1084	9	PTU89797
31	983.4	68.2	1059	9	AF052539
32	983.4	68.2	1071	6	AR119947
33	982.8	68.2	1060	9	AF011500
34	982.8	68.2	1060	9	AF011502
35	982.8	68.2	1060	9	AF011506
36	982.8	68.2	1060	9	AF011508
37	982.8	68.2	1060	9	AF011510
38	982.8	68.2	1060	9	AF011514
39	982.8	68.2	1060	9	AF011517
40	982.8	68.2	1060	9	AF011518
41	982.8	68.2	1060	9	AF011521
42	982.8	68.2	1060	9	AF011522
43	982.8	68.2	1060	9	AF011523
44	982.8	68.2	1060	9	AF011527
45	982.8	68.2	1060	9	AF011528

ALIGNMENTS

RESULT 1	AX427283	1442 bp	DNA	linear	PAT 18-JUN-2002
LOCUS	Sequence 5 from Patent EP199360.				
DEFINITION	AX427283				
ACCESSION	AX427283				
VERSION	AX427283.1	GI:21530636			
SOURCE	unidentified.				
ORGANISM	unidentified				
REFERENCE	1				
AUTHORS	Samson, M., Parmentier, M., Vassart, G. and Libert, F.				
TITLE	Active and inactive cc-chemokine receptors and nucleic acid				
JOURNAL	molecules encoding said receptor				
	Patent: EP 1199360-A 5 24-APR-2002;				

Euroscreen S.A. (BE)  
 Location/Qualifiers  
 1. .1442  
 /organism="unidentified"  
 /db\_xref="taxon:32644"  
 240. .887  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD35701.1"  
 /db\_xref="GI:21530637"  
 /translation="MDYQVSSPIYDINVTSEPCOKINVKOIAARLLPPLYSLVIFG  
 FVGNLVLILNKKRLKMTDIYLLNLAISDLFLFTVPFWAHVAOANDGNMCO  
 LTLGFLYFESGFERILLTLDRLVAVHVAFLKARYFEGVTVSYIWWVAFAS  
 LPIGIFRSQEGHUTCSHPYIKDHLGAGPAAACHGHLLENPKNSASVSK"  
 BASE COUNT 363 a 343 c 314 g 422 t  
 ORIGIN  
 Query Match 100.0%; Score 1442; DB 6; Length 1442;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GAATCCCCCAGCAGCCAGCTCCATCTAGTGACAGGGAAGCTAGCAGCAACCT 60  
 DB 1 GAATCCCCCAGCAGCCAGCTCCATCTAGTGACAGGGAAGCTAGCAGCAACCT 60  
 QY 61 TCCCTTCACCTACAAAACCTCATCTGGCCAAAAGAGAGTTAATCAATGTAGATC 120  
 DB 61 TCCCTTCACCTACAAAACCTCATCTGGCCAAAAGAGAGTTAATCAATGTAGATC 120  
 QY 121 TATGTAGGCAATTAACCACTATTGATGATATAACAGTTTGCAATTCATGGAGGCACT 180  
 DB 121 TATGTAGGCAATTAACCACTATTGATGATATAACAGTTTGCAATTCATGGAGGCACT 180  
 QY 181 AATACATCTAGGCTTTTAAAGATCACTTTTATTTATCCACAGGTTGGAACAGA 240  
 DB 181 AATACATCTAGGCTTTTAAAGATCACTTTTATTTATCCACAGGTTGGAACAGA 240  
 QY 241 TGGATTATCAAGTCAAGTCCATCTATGACATCAATTTATATCATCGGAGCCCTGCC 300  
 DB 241 TGGATTATCAAGTCAAGTCCATCTATGACATCAATTTATATCATCGGAGCCCTGCC 300  
 QY 301 AAAAATCAATGTGAGCAATCGAGCCCGCTCGCTCGCTCGCTCACTACCTGGTGT 360  
 DB 301 AAAAATCAATGTGAGCAATCGAGCCCGCTCGCTCGCTCGCTCACTACCTGGTGT 360  
 QY 361 TCATCTTTGGTTTGGGCAACATCGTCTCATCTCATCTCATCTCATCTCATCTCATCT 420  
 DB 361 TCATCTTTGGTTTGGGCAACATCGTCTCATCTCATCTCATCTCATCTCATCTCATCT 420  
 QY 421 TGAAGACATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCTTC 480  
 DB 421 TGAAGACATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCTTC 480  
 QY 481 TACTGTCCTCTTGGCTCACTATGTCGCCGCCAGTGGACITTTGGAATACATGT 540  
 DB 481 TACTGTCCTCTTGGCTCACTATGTCGCCGCCAGTGGACITTTGGAATACATGT 540  
 QY 541 GTCAACTCTTGACAGGCTCTATTTATAGGCTTCTCTGGAATCTTCTTCATATCC 600  
 DB 541 GTCAACTCTTGACAGGCTCTATTTATAGGCTTCTCTGGAATCTTCTTCATATCC 600  
 QY 601 TCCTGACAATCGATAGTACTGGCTGCTGCTCAATGCTGTGTTGTTTAAAGCCAGGA 660  
 DB 601 TCCTGACAATCGATAGTACTGGCTGCTGCTCAATGCTGTGTTGTTTAAAGCCAGGA 660  
 QY 661 CGGTCACTTTGGGTGTGACAGTGTATCATCTGGGTGGTGGTGGTGGTGGTGGTGGT 720  
 DB 661 CGGTCACTTTGGGTGTGACAGTGTATCATCTGGGTGGTGGTGGTGGTGGTGGTGGT 720  
 QY 721 TCCAGGAATCATCTTTACAGATCTCAAAAAGAGTCTTCAATACACCTGCAGCTCTC 780  
 DB 721 TCCAGGAATCATCTTTACAGATCTCAAAAAGAGTCTTCAATACACCTGCAGCTCTC 780

QY 781 ATTTCCATACATAAAGATAGTCACTTTGGGCTGGTCTCTGGCGTCTGCTGTCATGTC 840  
 DB 781 ATTTCCATACATAAAGATAGTCACTTTGGGCTGGTCTCTGGCGTCTGCTGTCATGTC 840  
 QY 841 ATCTGCTACTCGGGAATCTTAAACACTCTGCTTCGGTGTGAAATGAGAGAGGAC 900  
 DB 841 ATCTGCTACTCGGGAATCTTAAACACTCTGCTTCGGTGTGAAATGAGAGAGGAC 900  
 QY 901 AGGCTGTGAGGGTTATCTTCCACCATCATGATGTTTATTTCTCTCTGGCTCCCTAC 960  
 DB 901 AGGCTGTGAGGGTTATCTTCCACCATCATGATGTTTATTTCTCTCTGGCTCCCTAC 960  
 QY 961 ACATGTCCTTCTCTGACACCTTCCAGGAATCTTTGGCTGGAATATATGTCAGTAGC 1020  
 DB 961 ACATGTCCTTCTCTGACACCTTCCAGGAATCTTTGGCTGGAATATATGTCAGTAGC 1020  
 QY 1021 TCTAACAGTTGGACCAAGCTATGAGGTGACAGAGACTCTTGGGATGACGACATGCTGC 1080  
 DB 1021 TCTAACAGTTGGACCAAGCTATGAGGTGACAGAGACTCTTGGGATGACGACATGCTGC 1080  
 QY 1081 ATCAACCCATCATCTATGCTTTTGGGGAGAGATTTCAGAACTACCTTACTCTTC 1140  
 DB 1081 ATCAACCCATCATCTATGCTTTTGGGGAGAGATTTCAGAACTACCTTACTCTTC 1140  
 QY 1141 TTCCAAAAGCACATTCGCAAGCTTCTGCAAAATGCTGTTCTATTTTCCAGCAAGGCT 1200  
 DB 1141 TTCCAAAAGCACATTCGCAAGCTTCTGCAAAATGCTGTTCTATTTTCCAGCAAGGCT 1200  
 QY 1201 CCCGAGCGACCAAGCTCAGTTTACACCCGATCCACTGGGGAGCAGGAAATATCTCTGGC 1260  
 DB 1201 CCCGAGCGACCAAGCTCAGTTTACACCCGATCCACTGGGGAGCAGGAAATATCTCTGGC 1260  
 QY 1261 TTGTGACACGACTCAAGTGGGCTGGTGGGCTGGTGGGCTGGTGGGCTGGTGGT 1320  
 DB 1261 TTGTGACACGACTCAAGTGGGCTGGTGGGCTGGTGGGCTGGTGGGCTGGTGGT 1320  
 QY 1321 TCATACACAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGT 1380  
 DB 1321 TCATACACAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGT 1380  
 QY 1381 TAGGGTCTAAGATTCATTCATTTATTTGGCATCTGTTTAAAGTAGATAGATCCGAAT 1440  
 DB 1381 TAGGGTCTAAGATTCATTCATTTATTTGGCATCTGTTTAAAGTAGATAGATCCGAAT 1440  
 QY 1441 TC 1442  
 DB 1441 TC 1442

RESULT 2  
 AF009962  
 LOCUS  
 DEFINITION Homo sapiens CC-chemokine receptor (CCR-5) gene, delta-32 allele, complete cds.  
 ACCESSION AF009962  
 VERSION AF009962.1 GI:3243092  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 Tse, L., Ehrenberg, P. K., Chang, G. and Michael, N. L.  
 Genomic Organization and Functional Characterization of the Complete Transcription Unit for the Chemokine Receptor CCR-5, a Major Entry Co-Receptor for HIV-1  
 Unpublished  
 REFERENCE 1 (bases 1 to 7422)  
 Tse, L., Ehrenberg, P. K., Chang, G. and Michael, N. L.  
 Direct Submission  
 Submitted (23-JUN-1997) Div. of Retrovirology, Walter Reed Army Institute of Research, 13 Taft Court, Suite 200, Rockville, MD 20850, USA

FEATURES	Location/Qualifiers	Db	
source	1..7422		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="3"		
	/map="3p21"		
gene	1936..7204		
	/gene="CCR-5"		
	/allele="delta-32"		
TATA_signal	1936..1940		
	/gene="CCR-5"		
TATA_signal	1945..1949		
	/gene="CCR-5"		
mRNA	join(1976..2018,3923..7204)		
	/gene="CCR-5"		
	/product="CC-chemokine receptor"		
exon	1976..2018		
	/gene="CCR-5"		
intron	2019..3922		
	/gene="CCR-5"		
misc_feature	3025..3252		
	/gene="CCR-5"		
	/note="single strand coverage"		
exon	3923..7204		
	/gene="CCR-5"		
CDS	3934..4581		
	/gene="CCR-5"		
	/function="major entry co-receptor for HIV-1"		
	/note="mutant protein containing 215 amino acids (versus wild-type 352 amino acids) in which the C-terminal 31 amino acids are translated out of frame."		
	/codon_start=1		
	/product="CC-chemokine receptor"		
	/protein_id="AAC23944.1"		
	/db_xref="GI:3243093"		
	/translation="MDYQVSSPIYDINYYSEPCQKINVKQIAARLLPPLYSLVIFGFVGNMLVILILNCKRLKMSDIDYLLNLAIISDLFLLITVPFWAHYAAAOQDFGNMCO LLTGLYFPGFSGIFITILLITDRYLAHVHAFALKARTVFGVTVSVITWYVAFAS LPIIIFTSQEGHLVTCSSHPPIKDSHLGAPAAACHGILLGNPKNSVSR"		
polya_signal	7169..7174		
	/gene="CCR-5"		
polya_signal	7181..7186		
	/gene="CCR-5"		
polya_site	7205		
BASE COUNT	2126 a 1477 c 1667 g 2152 t		
ORIGIN			
Query Match	98.3%; Score 1417.2; DB 9; Length 7422;		
Best Local Similarity	99.6%; Pred. No. 0;		
Matches 1432; Conservative	0; Mismatches 3; Indels 3; Gaps 1;		
QY	1 GAATTCCTCCACAGAGCAAGCTCTCCATCTAGTGGACAGGGAAGCTAGCAGCAACCT	60	
Db	3695 GAATTCCTCCACAGAGCAAGCTCTCCATCTAGTGGACAGGGAAGCTAGCAGCAACCT	3754	
QY	61 TCCCTTCACTACAAACCTTCATGCTTGGCCAAAAGAGAGTTAATCAATGTAGACATC	120	
Db	3755 TCCCTTCACTACAAACCTTCATGCTTGGCCAAAAGAGAGTTAATCAATGTAGACATC	3814	
QY	121 TATGTAGCAATTAATAACCTATTGATGATATAAACAAGTTTGCATTCATGAGGGCAACT	180	
Db	3815 TATGTAGCAATTAATAACCTATTGATGATATAAACAAGTTTGCATTCATGAGGGCAACT	3874	
QY	181 AAATACATTTCTAGGACTTTATAAAGATCACATTTTATTTATGCACAGGGTGGACAGA	240	
Db	3875 AAATACATTTCTAGGACTTTATAAAGATCACATTTTATTTATGCACAGGGTGGACAGA	3934	
QY	241 TGGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATATTATATCATCGGAGCCCTGCC	300	
Db	3935 TGGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATATTATATCATCGGAGCCCTGCC	3994	
QY	301 AAAAATCAATGTGAAGCAATCGCAGCCGCTCTCGCTCTACTCACTGCTGT	360	
Db	5075 TTATAGAGGGTCTAAGATTATCCATTTATTTGGCATCTGTTTAAAGTAGATTAGATC	5132	
3995	AAAAATCAATGTGAAGCAATCGCAGCCGCTCTCGCTCTCGCTCTACTCACTGCTGT	4054	
361	TCATCTTTGGTTTTGTGGCAACATGCTGTGCATCTCTCATCTGATAAACTGCAAAAGGC	420	
4055	TCATCTTTGGTTTTGTGGCAACATGCTGTGCATCTCTCATCTGATAAACTGCAAAAGGC	4114	
421	TGAAGAGCATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTCTCTTC	480	
4115	TGAAGAGCATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTCTCTTC	4174	
481	TTACTGTCCCTTCTGGGCTCACTATCTGCCGCCACGTGGGACTTTGGAATACATGT	540	
4175	TTACTGTCCCTTCTGGGCTCACTATCTGCCGCCACGTGGGACTTTGGAATACATGT	4234	
541	GTCAACTCTTTGACAGGGCTCTATTTTATAGGCTTCTCTCTGGAATCTTCTTCATCATCC	600	
4235	GTCAACTCTTTGACAGGGCTCTATTTTATAGGCTTCTCTCTGGAATCTTCTTCATCATCC	4294	
601	TCCTGACAATCGATAGGTACCTGGCTGTCTCCATGCTGTGTTCGTTTAAAGCCAGGA	660	
4295	TCCTGACAATCGATAGGTACCTGGCTGTCTCCATGCTGTGTTCGTTTAAAGCCAGGA	4354	
661	CGGTCACTTTGGGCTGTGACAAAGTGTGATCACTTGGGTGGTGGCTGTGTTCGCTC	720	
4355	CGGTCACTTTGGGCTGTGACAAAGTGTGATCACTTGGGTGGTGGCTGTGTTCGCTC	4414	
721	TCCAGGAATCACTTTTACAGATCTCAAAAGAGGCTTTCATTAACCTCGAGCTC	780	
4415	TCCAGGAATCACTTTTACAGATCTCAAAAGAGGCTTTCATTAACCTCGAGCTC	4474	
781	ATTTTCCATCATTAAGAATAGTCACTCTGGGCTGGTCTCGCGCTGCTGTCTCATGTC	840	
4475	ATTTTCCATCATTAAGAATAGTCACTCTGGGCTGGTCTCGCGCTGCTGTCTCATGTC	4534	
841	ATCTCTACTCTGGGAACTCTAAAACCTCTCTTCGGTGTGGAATGAGAAGAGGCAC	900	
4535	ATCTCTACTCTGGGAACTCTAAAACCTCTCTTCGGTGTGGAATGAGAAGAGGCAC	4594	
901	AGGGCTGTGAGGCTTATCTTCCATCATGATTTGTTTATTTTCTCTCTGGGCTCCCTAC	960	
4595	AGGGCTGTGAGGCTTATCTTCCATCATGATTTGTTTATTTTCTCTCTGGGCTCCCTAC	4654	
961	AACATTTGCTCTCTCTGAACCTTCCAGGAATCTTTGGCTGGAATATGTCAGTAGC	1020	
4655	AACATTTGCTCTCTCTGAACCTTCCAGGAATCTTTGGCTGGAATATGTCAGTAGC	4714	
1021	TCTAACAGTTGGACCAAGCTATGACAGTGCAGAGACTCTTGGGATGACGCACTGCTGC	1080	
4715	TCTAACAGTTGGACCAAGCTATGACAGTGCAGAGACTCTTGGGATGACGCACTGCTGC	4774	
1081	ATCAACCCCATCATCTATGCTTGTGCGGAGAGTTTCAAGAACTACCTCTTAGCTTC	1140	
4775	ATCAACCCCATCATCTATGCTTGTGCGGAGAGTTTCAAGAACTACCTCTTAGCTTC	4834	
1141	TTCCAAAAGACATTCGCAACGCTTCTGCAATGCTGTCTTATTTTCCAGCAAGGCT	1200	
4835	TTCCAAAAGACATTCGCAACGCTTCTGCAATGCTGTCTTATTTTCCAGCAAGGCT	4894	
1201	CCGAGCGAGCAAGCTCAGTTTACACCCGATCCACATGCGGAGCAGGAAATATCTGTGGC	1260	
4895	CCGAGCGAGCAAGCTCAGTTTACACCCGATCCACATGCGGAGCAGGAAATATCTGTGGC	4954	
1261	TTGTACACAGGACTCAAGTGGGCTGACCCAGTCCAGTTGTGCACATGCTTAGTTT	1320	
4955	TTGTACACAGGACTCAAGTGGGCTGACCCAGTCCAGTTGTGCACATGCTTAGTTT	5014	
1321	TCATACACAGCTGGGCTGGGCTGG---TTGGGAGGCTTTTTTAAAGAGGTTACTG	1377	
5015	TCATACACAGCTGGGCTGGGCTGGGAGAGGCTTTTTTAAAGAGGTTACTG	5074	
1378	TTATAGAGGGTCTAAGATTATCCATTTATTTGGCACTGTTTAAAGTAGATTAGATC	1435	
5075	TTATAGAGGGTCTAAGATTATCCATTTATTTGGCATCTGTTTAAAGTAGATTAGATC	5132	

**RESULT 3**  
**AX153803**  
**LOCUS** AX153803 1477 bp DNA linear PAT 22-JUN-2001  
**DEFINITION** Sequence 1 from Patent EP1118858.  
**ACCESSION** AX153803  
**VERSION** AX153803.1 GI:14535433  
**KEYWORDS**  
**SOURCE** human.  
**ORGANISM** Homo sapiens  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**REFERENCE** 1 (bases 1 to 1477)  
**AUTHORS** Dobbs, S., Perros, M. and Rickett, G.A.  
**TITLE** Assay method  
**JOURNAL** Patent: EP 1118858-A 1 25-JUL-2001;  
 Pfizer Limited (GB); PFIZER INC. (US)  
**FEATURES** Location/Qualifiers  
 1..1477  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
**BASE COUNT** 374 a 349 c 320 g 431 t 3 others  
**ORIGIN**  
 Query Match 95.9%; Score 1382.8; DB 6; Length 1477;  
 Best Local Similarity 97.4%; Pred. No. 0;  
 Matches 1439; Conservative 0; Mismatches 3; Indels 35; Gaps 2;  
 QY 1 GAATTCGCCCAACAGAGCCAGCTCTCCATCTAGTGGACAGGAAGTAGCAGCAAACT 60  
 DB 1 GAATTCGCCCAACAGAGCCAGCTCTCCATCTAGTGGACAGGAAGTAGCAGCAAACT 60  
 QY 61 TCCCTTCACTACAAATTCATTGCTTGGCCAAAAGAGAGTTAATTCATGTAGACATC 120  
 DB 61 TCCCTTCACTACAAATTCATTGCTTGGCCAAAAGAGAGTTAATTCATGTAGACATC 120  
 QY 121 TATGTAGGCAATTAATAAAGCTATGATGTATATAAAGAGTTGCAATTCATGAGGGCAACT 180  
 DB 121 TATGTAGGCAATTAATAAAGCTATGATGTATATAAAGAGTTGCAATTCATGAGGGCAACT 180  
 QY 181 AATACATCTAGAGCTTTATAAAGATCACTTTTATTATGACAGAGGTGGAACAAGA 240  
 DB 181 AATACATCTAGAGCTTTATAAAGATCACTTTTATTATGACAGAGGTGGAACAAGA 240  
 QY 241 TGGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATTTATATACATCGGAGCCCTGCC 300  
 DB 241 TGGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATTTATATACATCGGAGCCCTGCC 300  
 QY 301 AAAAATCAATGTGAAGCAAAATCGAGCCGCCCTCTCGCTCCGCTCTACTCACTGGTGT 360  
 DB 301 AAAAATCAATGTGAAGCAAAATCGAGCCGCCCTCTCGCTCCGCTCTACTCACTGGTGT 360  
 QY 361 TCATCTTTGGTTTTGGGCAACATGCTGTATCTCACTCACTCACTCACTCACTCACTCACT 420  
 DB 361 TCATCTTTGGTTTTGGGCAACATGCTGTATCTCACTCACTCACTCACTCACTCACTCACT 420  
 QY 421 TGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
 DB 421 TGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480  
 QY 481 TTACTGTCCCTTCTGGGCTCACTATGCTGCGCCGAGTGGGACTTTTGGAAATACAATGT 540  
 DB 481 TTACTGTCCCTTCTGGGCTCACTATGCTGCGCCGAGTGGGACTTTTGGAAATACAATGT 540  
 QY 541 GTCACTCTTGACAGGCTCTATTTTATAGGCTTCTCTGGAATCTTCTTCATCATCC 600  
 DB 541 GTCACTCTTGACAGGCTCTATTTTATAGGCTTCTCTGGAATCTTCTTCATCATCC 600  
 QY 601 TCTTGACATCGATAGTACCTGGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
 DB 601 TCTTGACATCGATAGTACCTGGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660

QY 561 CGGTACACCTTTGGGGTGGTGACAAAGTGTGATCACTTTGGGTGGTGGTGGTGGTGGTGGTGGT 720  
 DB 561 CGGTACACCTTTGGGGTGGTGACAAAGTGTGATCACTTTGGGTGGTGGTGGTGGTGGTGGTGGT 720  
 QY 721 TCCAGAGATCATCTTTACAGATCTCAAAAAGAGTCTTTCATTACACCTGAGCTCTC 780  
 DB 721 TCCAGAGATCATCTTTACAGATCTCAAAAAGAGTCTTTCATTACACCTGAGCTCTC 780  
 QY 781 ATTTTCCAT-----ACATTAAAGATAGTCACT 808  
 DB 781 ATTTTCCATAGCTAGTCACTATCAATCTGAGAGATTTCCAGACATTAAGATAGTCACT 840  
 QY 809 TGGGGTGGTCTCCGCTGCTTGTTCATGTCATCTGCTACTCGGGAATCCTAAAAATC 868  
 DB 841 TGGGGTGGTCTCCGCTGCTTGTTCATGTCATCTGCTACTCGGGAATCCTAAAAATC 900  
 QY 869 TGGTTCGGTGTGAAATGAGAAGAGGACAGGCTGTGAGGCTTATCTTCACCATCA 928  
 DB 901 TGGTTCGGTGTGAAATGAGAAGAGGACAGGCTGTGAGGCTTATCTTCACCATCA 960  
 QY 929 TGATTGTTTATTTCTTCTTGGGCTCCCTACAAATTTCTTCTTCTGAAACATCTCC 988  
 DB 961 TGATTGTTTATTTCTTCTTGGGCTCCCTACAAATTTCTTCTTCTGAAACATCTCC 1020  
 QY 989 AGGAATCTTTGGGCTCAATTAATTCAGTGTCTTACAGGTTGGACCAAGCTATGCAAG 1048  
 DB 1021 AGGAATCTTTGGGCTCAATTAATTCAGTGTCTTACAGGTTGGACCAAGCTATGCAAG 1080  
 QY 1049 TGACAGAGACTCTTTGGGATGACGCACTGCTCATCAACCCCATCATCTATGCTTTGTCG 1108  
 DB 1081 TGACAGAGACTCTTTGGGATGACGCACTGCTCATCAACCCCATCATCTATGCTTTGTCG 1140  
 QY 1109 GGGAGAGTTCAGAACTACCTCTTAGTCTTCTTCCAAAGACATTTGCCAAACGCTCT 1168  
 DB 1141 GGGAGAGTTCAGAACTACCTCTTAGTCTTCTTCCAAAGACATTTGCCAAACGCTCT 1200  
 QY 1169 GCAATGCTGTCTTATTTTCCAGAGAGGCTCCGAGGAGCAAGCTCAAGTTCAGTTCACACC 1228  
 DB 1201 GCAATGCTGTCTTATTTTCCAGAGAGGCTCCGAGGAGCAAGCTCAAGTTCAGTTCACACC 1260  
 QY 1229 GATCCACTGGGAGCAGGAATAATCTGTGGCTTTGTGACACGACTCAAGTGGGCTGGTG 1288  
 DB 1261 GATCCACTGGGAGCAGGAATAATCTGTGGCTTTGTGACACGACTCAAGTGGGCTGGTG 1320  
 QY 1289 ACCAGTCAAGTGTGCACATGGCTTAGTTCATACACAGCTTGGGCTGGGCTGGGCTGGG 1348  
 DB 1321 ACCAGTCAAGTGTGCACATGGCTTAGTTCATACACAGCTTGGGCTGGGCTGGGCTGGG 1380  
 QY 1349 GG---GAGTCTTTTAAAGGAAGTTACTGTTATAGAGGCTCTAAGATTTCATCATTT 1405  
 DB 1381 TGGNAGGCTCTTTTAAAGGAAGTTACTGTTATAGAGGCTCTAAGATTTCATCATTT 1440  
 QY 1406 ATTGTCATCTGTTTAAAGTAGATTAGATCGGAATTC 1442  
 DB 1441 ATTGTCATCTGTTTAAAGTAGATTAGATCGGAATTC 1477

**RESULT 4**  
**AX427281**  
**LOCUS** AX427281 1477 bp DNA linear PAT 18-JUN-2002  
**DEFINITION** Sequence 3 from Patent EP1199360.  
**ACCESSION** AX427281  
**VERSION** AX427281.1 GI:21530634  
**KEYWORDS**  
**SOURCE** unidentified.  
**ORGANISM** unidentified.  
**REFERENCE** 1  
**AUTHORS** Samson, M., Parmentier, M., Vassart, G. and Libert, F.  
**TITLE** Active and inactive cc-chemokine receptors and nucleic acid  
 molecules encoding said receptor  
**JOURNAL** Patent: EP 1199360-A 3 24-APR-2002;  
 Euroscreen S.A. (BE)

QY	1	GAATTCCCCCAACAGAGCAAGCTCTCCATCTACTGTGACAGGGAAGCTAGCAGCAAAAGCT	60
Db	1	GAATTCCCCCAACAGAGCAAGCTCTCCATCTACTGTGACAGGGAAGCTAGCAGCAAAAGCT	60
QY	61	TCCCTTCACTCAAAAACCTTCATGCTGTGGCCAAAAGAGAGTAAATCAATGTAGACATC	120
Db	61	TCCCTTCACTCAAAAACCTTCATGCTGTGGCCAAAAGAGAGTAAATCAATGTAGACATC	120
QY	121	TATGTAGGCAATTAAAAACCTATTGATGTATAAACAGTTGCATTTCATGAGGGCAACT	180
Db	121	TATGTAGGCAATTAAAAACCTATTGATGTATAAACAGTTTGCATTTCATGAGGGCAACT	180
QY	181	AAATACATCTAGGACTTTATAAAGATCACCTTTTATTATGACAGGGTGGAAACAGA	240
Db	181	AAATACATCTAGGACTTTATAAAGATCACCTTTTATTATGACAGGGTGGAAACAGA	240
QY	241	TGSAATTATCAAGTGTCAAGTCCAACTATGACATCAATATTATACATCGAGGCCCTGCC	300
Db	241	TGSAATTATCAAGTGTCAAGTCCAACTATGACATCAATATTATACATCGAGGCCCTGCC	300
QY	301	AAAAAATCAATGTGAAGCAAAATCGCAGCCCGCTCTGGCTCTACTCACTGGTGT	360
Db	301	AAAAAATCAATGTGAAGCAAAATCGCAGCCCGCTCTGGCTCTACTCACTGGTGT	360
QY	361	TCATCTTTGGTTTGTGGGCAACATGCTGGTCACTCCTCACTGATAAATGCAAAAGGC	420
Db	361	TCATCTTTGGTTTGTGGGCAACATGCTGGTCACTCCTCATCTGATAAATGCAAAAGGC	420
QY	421	TGAAGAGATGACTGACATCTACCATCTGCTCAACCTGGCCATCTCTGACCTGTTTTCTCTC	480
Db	421	TGAAGAGATGACTGACATCTACCATCTGCTCAACCTGGCCATCTCTGACCTGTTTTCTCTC	480
QY	481	TTACTGTCCCTCTCTGGGCTCACTATGCTGGCCGCCAGTGGGACTTTGGAAAATACAAATG	540
Db	481	TTACTGTCCCTCTCTGGGCTCACTATGCTGGCCGCCAGTGGGACTTTGGAAAATACAAATG	540
QY	541	GTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTCTCTGGAATCTTCTTCATCATCC	600
Db	541	GTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTCTCTGGAATCTTCTTCATCATCC	600
QY	601	TCCTGACAATCGATAGGTACCTGGCTGTGGTCCATGCTGTGTGTTTTAAAGCCAGGA	660
Db	601	TCCTGACAATCGATAGGTACCTGGCTGTGGTCCATGCTGTGTGTTTTAAAGCCAGGA	660
QY	661	CGGTCACTTTGGGTTGTGAAGTGTGATCACTTGGGTGGTGGCTGTGTGTTGGCTCTC	720
Db	661	CGGTCACTTTGGGTTGTGAAGTGTGATCACTTGGGTGGTGGCTGTGTGTTGGCTCTC	720
QY	721	TCCAGGAATCATCTTTACCAAGATCTCAAAAAGAGTCTTCATTACACCTTGACGCTCTC	780

TITLE		Direct Submission	
JOURNAL	Submitted (23-OCT-1997) Medicine, University of Texas Health Science Center at San Antonio, 7703, Floyd Curl Drive, San Antonio, TX 78284, USA		
FEATURES		Location/Qualifiers	
source	1. .6059 /organism="Homo sapiens" /db_xref="taxon:9606" Join(AF031236.1:1..1976,1..6059) /gene="CCR5" Join(1..57,559..847,2751..6059) /gene="CCR5" /product="CC chemokine receptor 5A" Join(1..57,794..847,2751..6059) /gene="CCR5" /product="CC chemokine receptor 5B" Join(750..847,2751..6059) /gene="CCR5" /product="CCR5 truncated isoform" Join(773..847,2751..6059) /gene="CCR5" /product="CCR5 truncated isoform" Join(777..847,2751..6059) /gene="CCR5" /product="CCR5 truncated isoform" Join(779..847,2751..6059) /gene="CCR5" /product="CCR5 truncated isoform" Join(784..847,2751..6059) /gene="CCR5" /product="CCR5 truncated isoform" Join(795..847,2751..6059) /gene="CCR5" /product="CCR5 truncated isoform" Join(798..847,2751..6059) /gene="CCR5" /product="CCR5 truncated isoform" Join(804..847,2751..6059) /gene="CCR5" /product="CCR5 truncated isoform" Join(805..847,2751..6059) /gene="CCR5" /product="CCR5 truncated isoform" Join(806..847,2751..6059) /gene="CCR5" /product="CCR5 truncated isoform" 2762..3820 /gene="CCR5" /codon_start=1 /product="CC chemokine receptor 5" /protein_id="AAB94735.1" /db_xref="GI:2739499" /translation="MDYQVSSPDIDNYITSPKQINVKQIAARLLPLXSLVFIFG FVGNMVLILLNKRKLSMDIYLNLAILDLPELLTVPFAHYAAQWDFGNTMQ LILGLIFPFGSIFPIILLIDRYLAVHYAFALKARTVTEGWVSVLTWVAVPAS LPGLIIFRSKGLHYTCSSFFPSYQFQWKFQILKIVILGLVLELLVWICISGLI KTLRLCRNKRHRVRLIFMIFIMVFLNVLINLLNTFFQFELNNCCSSNRID QAMQVITLGMTHCCINPIIYAFVGEKFRNLLVFFQHKIAKRFCKCSIFQEAPEP ASSYITSTGEQIEISVGL"		
BASE COUNT	1753 a 1210 c 1393 g 1703 t		
ORIGIN			
Query Match 95.4%; Score 1375.2; DB 9; Length 6059; Best Local Similarity 97.4%; Pred. No. 0; Matches 1432; Conservative 0; Mismatches 3; Indels 35; Gaps 2;			
QY	1 GAATTCCTCCCAACAGCCAAAGCTCTCCATCTAGTCGACAGGAGAGCTAGCAGCAAACT 60 		
Db	2523 GAATTCCTCCCAACAGCCAAAGCTCTCCATCTAGTCGACAGGAGAGCTAGCAGCAAACT 2582 		
QY	61 TCCCTTCACTACAAACTTCATGCTTGGCCAAAGAGAGAGTAAATTCATGTAGACATC 120 		
Db	2583 TCCTTCACTACAAACTTCATGCTTGGCCAAAGAGAGAGTAAATTCATGTAGACATC 2642 		

QY	1169	GCAATAGCTGTTTCATTTTCCAGCAAGAGGCTCCGAGCGAGCAAGCTCAGTTTACACC	1228
Db	3723	GCAATAGCTGTTTCATTTTCCAGCAAGAGGCTCCGAGCGAGCAAGCTCAGTTTACACC	3782
QY	1229	GATCCACTGGGAGCAGGAAATATCTGTGGGCTTGTCACAGGACTCAAGTGGCGTGGT	1288
Db	3783	GATCCACTGGGAGCAGGAAATATCTGTGGGCTTGTCACAGGACTCAAGTGGCGTGGT	3842
QY	1289	ACCAGTCAGAGTTGTGCACATGGCTTAGTTTTCATACAGACCTGGGCTGGGGTGG--	1346
Db	3843	ACCAGTCAGAGTTGTGCACATGGCTTAGTTTTCATACAGACCTGGGCTGGGGTGGGG	3902
QY	1347	TTGGAGGCTCTTTTAAAGGAGGACTTACTGTTATAGAGGCTTATAGAGGCTTACATCCATT	1405
Db	3903	TGGGAGAGGCTCTTTTAAAGGAGGACTTACTGTTATAGAGGCTTATAGAGGCTTACATCCATT	3962
QY	1406	ATTGGCATCTGTTTAAAGTAGATTAGATC	1435
Db	3963	ATTGGCATCTGTTTAAAGTAGATTAGATC	3992
RESULT 6			
LOCUS	AX335952	143068 bp	DNA
DEFINITION	Sequence 6461 from Patent WO0194629.		linear
ACCESSION	AX335952		
VERSION	AX335952.1	GI:18126671	
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Young, P.E., Augustus M., Carter, K.C., Ebner, R., Endress, G., Horrigan, S., Soppet, D.R. and Weaver, Z.		
TITLE	Cancer gene determination and therapeutic screening using signature gene sets		
JOURNAL	Patent: WO 0194629-A 6461 13-DEC-2001;		
FEATURES	Avalon Pharmaceuticals (US)		
SOURCE	Location/Qualifiers		
	1..143068		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	41194 a 30122 c 32403 g 39349 t		
ORIGIN			
Query Match	95.4%;	Score 1375.2;	DB 6; Length 143068;
Best Local Similarity	97.4%;	Pred. No. 0;	
Matches 1432;	Conservative 0;	Mismatches 3;	Indels 35; Gaps 2;
QY	1	GAATTCCTCCCAACAGAGCCAAAGCTCTCCATCTAGTGAGCAGGGAAGCTAGCAGCAAACT	60
Db	61244	GAATTCCTCCCAACAGAGCCAAAGCTCTCCATCTAGTGAGCAGGGAAGCTAGCAGCAAACT	61303
QY	61	TCGCTTCACACAAACTTCATTTGCTTGGCCCAAAAGAGAGTTAATTCATGTTAGACATC	120
Db	61304	TCCTTCACACAAACTTCATTTGCTTGGCCCAAAAGAGAGTTAATTCATGTTAGACATC	61363
QY	121	TATGTAGGCAATTA AAAACCTATTGATGTATAAACAGTTTCATTTCATGGAGGCAACT	180
Db	61364	TATGTAGGCAATTA AAAACCTATTGATGTATAAACAGTTTCATTTCATGGAGGCAACT	61423
QY	181	AAATACATCTTAGGACTTTATAAAGATCACTTTTATTATGTCACAGGTTGGAAACAAGA	240
Db	61424	AAATACATCTTAGGACTTTATAAAGATCACTTTTATTATGTCACAGGTTGGAAACAAGA	61483
QY	241	TGGATTATCAAGTGTCAAGTCCAACTCATGTACATCAATATTATACATCGAGCCCTGCC	300
Db	61484	TGGATTATCAAGTGTCAAGTCCAACTCATGTACATCAATATTATACATCGAGCCCTGCC	61543
QY	301	AAAAATCAATGTGAGCAAAATCGAGCCGCCCTTCCTGGCTCCGCTCTACTCACTGGTGT	360
Db	61544	AAAAATCAATGTGAGCAAAATCGAGCCGCCCTTCCTGGCTCCGCTCTACTCACTGGTGT	61603

QY 1406 ATTGGCATCTGTTTAAGTAGATTAGATC 1435  
 |||||  
 Db 62684 ATTGGCATCTGTTTAAGTAGATTAGATC 62713

RESULT 7  
 HSU95626 143068 bp DNA linear PRI 16-MAY-1997  
 LOCUS Homo sapiens ccr2b (ccr2), ccr2a (ccr2), ccr5 (ccr5) and ccr6  
 DEFINITION (ccr6) genes, complete cds, and lactoferrin (lactoferrin) gene,  
 partial cds, complete sequence.  
 ACCESSION U95626  
 VERSION U95626.1 GI:2104517  
 KEYWORDS HTG.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 143068)  
 McCombie, W.R., Wilson, R., Chen, E., Gibbs, R., Zuo, L., Johnson, D.,  
 Nhan, M., Parnell, L., Dedhia, N., Ansari, A., Mardis, E., Schutz, K.,  
 Gnoj, L., de la Bastide, M., Kaplan, N., Greco, T., Touchman, J.,  
 Muzny, D., Chen, C.-N., Evans, C., Fitzgerald, M., See, L.H., Tang, M.,  
 Porcel, B.M., Dragan, Y., Giacalone, J., Pae, A., Powell, E.,  
 Solinsky, K.A., Desilva, U., Diaz-Perez, S., Zhou, X., Yu, Y.,  
 Watanabe, M., Doggett, N., Garcia, D. and Sagripanti, J.-L.  
 Human BAC clone 110P12  
 Unpublished (1997)  
 2 (bases 1 to 143068)  
 McCombie, R.W., Wilson, R., Chen, E., Gibbs, R., Zuo, L., Johnson, D.,  
 Nhan, M., Parnell, L., Dedhia, N., Ansari, A., Mardis, E., Schutz, K.,  
 Gnoj, L., de la Bastide, M., Kaplan, N., Greco, T., Touchman, J.,  
 Muzny, D., Chen, C.-N., Evans, C., Fitzgerald, M., See, L.H., Tang, M.,  
 Porcel, B.M., Dragan, Y., Giacalone, J., Pae, A., Powell, E.,  
 Solinsky, K.A., Desilva, U., Diaz-Perez, S., Zhou, X., Yu, Y.,  
 Watanabe, M., Doggett, N., Garcia, D. and Sagripanti, J.-L.  
 Direct Submission  
 TITLE Submitted (27-MAR-1997) Advanced Genome Sequence Analysis Course,  
 Cold Spring Harbor Laboratory, 1Bungtown Rd., Cold Spring Harbor,  
 NY 11724, USA  
 COMMENT Regions with single-strand coverage are as follows:

31434 - 31443 37900 - 37968 53303 - 53357  
 59166 - 59206 63708 - 63998 65200 - 65335  
 78605 - 78713 92135 - 92137 112377 - 112551  
 112643 - 112778 134284 - 134309 134914 - 135019  
 143046 - 144068.

FEATURES  
 source  
 1. .143068  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="3"  
 /clone="BAC 110P12"  
 46056..49505  
 /gene="ccr2"  
 /note="confirmed by similarity to Human monocyte  
 chemoattractant protein 1 receptor (ccr2) mRNA (Accession  
 Number U80924), two alternatively spliced mRNAs."  
 join(46056..47046,48255..49505)  
 /gene="ccr2"  
 /product="ccr2a"  
 /note="confirmed by similarity to Human monocyte  
 chemoattractant protein 1 receptor (ccr2) alternatively  
 spliced mRNA encoding A-form carboxyl tail, Accession  
 Number U80924."  
 46056..47997  
 /gene="ccr2"  
 /product="ccr2b"  
 /note="confirmed by similarity to Human monocyte  
 chemoattractant protein 1 receptor (ccr2) alternatively  
 spliced mRNA encoding B-form carboxyl tail. Accession  
 Number: U80924."  
 join(46106..47046,48255..48438)  
 CDS

/gene="ccr2"  
 /note="confirmed by similarity to Human monocyte  
 chemoattractant protein 1 receptor (ccr2) alternatively  
 spliced A-form, Encoded by GenBank Accession Number  
 U80924, gi 1168965"  
 /codon\_start=1  
 /product="ccr2a"  
 /protein\_id="AA857791.1"  
 /db\_xref="GI:2104518"  
 /translation="MLSTSRSRFIRNTNESGEVTFDDYDGAPCHKFDVKQIGAL  
 LPPLYSIVTFPGVGNMVLVLLINCKKLCITDIYLAIVHAVFALKARTVFGVVT  
 ANEWFGNACMKTGLYHIGYFGGFIITLLIDRYLAIVHAVFALKARTVFGVVT  
 SVTLVAVASVPGIITFKCKEDSVVCGVPEPGWNNFTIMRNLGLVPLILM  
 VICYSILATLLRCNEKKRRAVRVITIMVYFLFWPTVYINVLANTQEEFGLSN  
 CESTSOLDQATQVETLGMTHCCINPIIYAFVGEKFRSLFHAGCTRIAPLQKPVCGG  
 PGVPRKNYKVTQGLLDGRGKSGRAPEASLQKEGA"  
 46106..47188  
 /gene="ccr2"  
 /note="confirmed by similarity to Human monocyte  
 chemoattractant protein 1 receptor (ccr2), Accession  
 Number: 1168965"  
 /codon\_start=1  
 /product="ccr2b"  
 /protein\_id="AA857792.1"  
 /db\_xref="GI:2104519"  
 /translation="MLSTSRSRFIRNTNESGEVTFDDYDGAPCHKFDVKQIGAL  
 LPPLYSIVTFPGVGNMVLVLLINCKKLCITDIYLAIVHAVFALKARTVFGVVT  
 ANEWFGNACMKTGLYHIGYFGGFIITLLIDRYLAIVHAVFALKARTVFGVVT  
 SVTLVAVASVPGIITFKCKEDSVVCGVPEPGWNNFTIMRNLGLVPLILM  
 VICYSILATLLRCNEKKRRAVRVITIMVYFLFWPTVYINVLANTQEEFGLSN  
 CESTSOLDQATQVETLGMTHCCINPIIYAFVGEKFRSLFHAGCTRIAPLQKCPV  
 FYRETVDGVTNTSTPSTGEQESAGL"  
 59531..64785  
 /gene="ccr5"  
 join(59531..59573,61472..64785)  
 /gene="ccr5"  
 /product="ccr5"  
 /note="confirmed by similarity to Human cc chemokine  
 receptor 5 (ccr5) mRNA. Accession number: U54994."  
 61483..62541  
 /gene="ccr5"  
 /note="confirmed by similarity to human CC chemokine  
 receptor 5 (ccr5) protein, encoded by GenBank Accession  
 Number U54994, gi 1457946"  
 /codon\_start=1  
 /product="ccr5"  
 /protein\_id="AA857793.1"  
 /db\_xref="GI:2104520"  
 /translation="MDYQVSPYDNYTSPQCQKINVKQIAARLLPLYSIVTFIFG  
 FVGNMVLILINCKRKMSTDIYLLNLAISDLFLITVFWAHVAAAQDFGNTMCQ  
 LLTGLYIFGFGFIITLLIDRYLAIVHAVFALKARTVFGVTSVITVWAVAFS  
 LPIGIFTRQKGLHYTCSSHPFSQYQFNKFTLVILGLVPLVWVICYSGIL  
 KTLRCNEKKRRAVRVITIMVYFLFWPTVYINVLANTQEEFGLSNSSRLD  
 QAMQVETLGMTHCCINPIIYAFVGEKFRSLFHAGCTRIAPLQKCPV  
 ASSVYTRSGEQEISVGL"  
 96634..97683  
 /product="ccr6"  
 /note="Putative mRNA identified by homology to CCR5 mRNA."  
 /evidence=not\_experimental  
 96642..97676  
 /gene="ccr6"  
 96642..97676  
 /gene="ccr6"  
 /note="Identified as a gene by Grail Version 1.3c.  
 Translated sequence exhibits 42% sequence identity to CCR5  
 protein."  
 /codon\_start=1  
 /evidence=not\_experimental  
 /product="ccr6"  
 /protein\_id="AA857794.1"  
 /db\_xref="GI:2104521"  
 /translation="NANYTAPDEYDVLIEGLESDEARCCCKYDAALSAQLVPSL  
 CSAVFVIGLDNLVLLVLLVKKLRVENIYLLNLAISDLFLITVFWAHVAAAQDFGNTMCQ



gene	CKILGLIFVLGYSETFNCLLTVORLVLHKGNEFFSARRRVPCCGLIISVLAWTAI LATLEPYVYKQPMQEDKYKAFSRTPFPADTFWKHFLFLKNNISVLPLFIPTF LVYQMKTLRFRRQYSLFKLVFAIMVFLMAYPIAFLSTFKHFLSLDCKSSY NLDKSVHITKLIAITHCICINPLLYAFIDGTFSKYLCRCFHLRSNTPLQPRGSAQGS REPDHSTEV
mRNA	complement(124605..>139255) /gene="lactoferrin" complement(join(124605..124816,126528..126717, 127884..128068,130006..130073,132023..132164, 133863..134018,135022..135075,135890..135980, 137445..137599,138436..138610,139077..>139255)) /gene="lactoferrin" /product="lactoferrin"
CDS	/note="confirmed by similarity to lactoferrin mRNA, accession number M73700" complement(join(124782..124816,126528..126717, 127884..128068,130006..130073,132023..132164, 133863..134018,135022..135075,135890..135980, 137445..137599,138436..138610,139077..>139253)) /gene="lactoferrin" /note="confirmed by similarity to lactoferrin protein, encoded by GenBank Accession Number M73700, gi 186818" /codon_start=1 /product="lactoferrin" /protein_id="AAB57795.1" /db_xref="GI:2104522"
BASE COUNT	41194 a 30122 c 32403 g 39349 t
ORIGIN	Query Match 95.4%; Score 1375.2; DB 9; Length 143068; Best Local Similarity 97.4%; Pred. No. 0; Matches 1432; Conservative 0; Mismatches 3; Indels 35; Gaps 2;
QY	1 GAATTCCTCCCAACAGCCAGCTCCATCTAGTGGACAGGAGCTAGCAGCAACCT 60
Db	61244 GAATTCCTCCCAACAGCCAGCTCCATCTAGTGGACAGGAGCTAGCAGCAACCT 61303
QY	61 TCCCTTCACATAAACTTCATGTTGGCCAAAAGAGAGTTAATCAATGTAGACATC 120
Db	61304 TCCCTTCACATAAACTTCATGTTGGCCAAAAGAGAGTTAATCAATGTAGACATC 61363
QY	121 TATGTAGCAATTAACAACTTATGATGTATTAACACAGTTGCAATTCATGGAGGCACT 180
Db	61364 TATGTAGCAATTAACAACTTATGATGTATTAACACAGTTGCAATTCATGGAGGCACT 61423
QY	181 AAATACATCTTAGGACTTTATAAAGATCACATTTTATTTATGCACAGGTTGGACACAGA 240
Db	61424 AAATACATCTTAGGACTTTATAAAGATCACATTTTATTTATGCACAGGTTGGACACAGA 61483
QY	241 TGGATATCACTGTCAGTCCCAATCTATGACATCAATTAATATACATCGGAGCCCTGCC 300
Db	61484 TGGATATCACTGTCAGTCCCAATCTATGACATCAATTAATATACATCGGAGCCCTGCC 61543
QY	301 AAAAAATCAATGTGAAGCAATCGACGCCGCTCCCTCCCTCCGCTCTACTCACTGGTGT 360
Db	61544 AAAAAATCAATGTGAAGCAATCGACGCCGCTCCCTCCCTCCGCTCTACTCACTGGTGT 61603
QY	361 TCATCTTTGGTTTGGGCAACATGCTGTCATCTCATCTGATCAATGATAAAGGCAAGGC 420
Db	61604 TCATCTTTGGTTTGGGCAACATGCTGTCATCTCATCTGATCAATGATAAAGGCAAGGC 61663
QY	421 TGAAGACATGACTGACATCTAOCCTCAACCTGGGCACTCTGACCTGTTTTCCTTC 480
Db	61664 TGAAGACATGACTGACATCTAOCCTCAACCTGGGCACTCTGACCTGTTTTCCTTC 61723

QY	481 TTACTGTCCCTCTCTGGGGTCACTATCTGCGCCCACTGGGACTTTGGAATAACAATGT 540
Db	61724 TTACTGTCCCTCTCTGGGGTCACTATCTGCGCCCACTGGGACTTTGGAATAACAATGT 61783
QY	541 GTCACTCTTTGACAGGGCTCTATTTATAGGCTCTCTCTGGAATCTTCTTCATCATCC 600
Db	61784 GTCACTCTTTGACAGGGCTCTATTTATAGGCTCTCTCTGGAATCTTCTTCATCATCC 61843
QY	601 TCCTGACAATCGATAGTAGTCTGCTGCTCCATCTGCTGCTTTTAAAGCCAGGA 660
Db	61844 TCCTGACAATCGATAGTAGTCTGCTGCTCCATCTGCTGCTTTTAAAGCCAGGA 61903
QY	661 CGGTCACTTTGGGGTGGTGAAGTGTGATCAGTCTGGGGTGGTGGTCTGCTCTC 720
Db	61904 CGGTCACTTTGGGGTGGTGAAGTGTGATCAGTCTGGGGTGGTGGTCTGCTCTC 61963
QY	721 TCCAGGAATCANTTTACAGATCTCAAAAGAGGCTTTTCAATACCTGACGCTCTC 780
Db	61964 TCCAGGAATCANTTTACAGATCTCAAAAGAGGCTTTTCAATACCTGACGCTCTC 62023
QY	781 ATTTTCCAT-----ACATTAAGATAGTCACT 808
Db	62024 ATTTTCCATACAGTCACTATCAATTTCTGGAAGATTTCCAGACATTAAGATAGTCACT 62083
QY	809 TGGGCTGGTCTCGCGCTGCTGCTGATGTCATCTGCTACTCGGAATCTTAAACTC 868
Db	62084 TGGGCTGGTCTCGCGCTGCTGCTGATGTCATCTGCTACTCGGAATCTTAAACTC 62143
QY	869 TCGTTCGGTGTGCAATGAGAAGAGGACACAGGCTGTGAGGCTTATCTTCACTCA 928
Db	62144 TCGTTCGGTGTGCAATGAGAAGAGGACACAGGCTGTGAGGCTTATCTTCACTCA 62203
QY	929 TGATTTGTTTATTTCTTCTTGGGCTCCCTTACACATTTGCTTCTTCTGGAACCTTCC 988
Db	62204 TGATTTGTTTATTTCTTCTTGGGCTCCCTTACACATTTGCTTCTTCTGGAACCTTCC 62263
QY	989 AGAATTTCTTGGCTGATATTTGCACTAGCTTAAACAGGTTGGACCAAGCTATGCAGG 1048
Db	62264 AGAATTTCTTGGCTGATATTTGCACTAGCTTAAACAGGTTGGACCAAGCTATGCAGG 62323
QY	1049 TGACAGAGCTCTTGGGAGGACGCTGCTGATCAACCCATCATCTATGCTTTTGTG 1108
Db	62324 TGACAGAGCTCTTGGGAGGACGCTGCTGATCAACCCATCATCTATGCTTTTGTG 62383
QY	1109 GGGAGAGTTTCAAGAACTACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1168
Db	62384 GGGAGAGTTTCAAGAACTACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 62443
QY	1169 GCAATGCTGTTCTTATTTTCCAGCAAGGCTCCGAGGAGCAAGCTCAGTTTACACCC 1228
Db	62444 GCAATGCTGTTCTTATTTTCCAGCAAGGCTCCGAGGAGCAAGCTCAGTTTACACCC 62503
QY	1229 GATCCACTGGGAGGACGAGAAATATCTTGGGCTTGTGACACGAGCTCAAGTGGGCTGTG 1288
Db	62504 GATCCACTGGGAGGACGAGAAATATCTTGGGCTTGTGACACGAGCTCAAGTGGGCTGTG 62563
QY	1289 ACCAGTCAAGTGTGACATGCTTGTGTTTATACACAGCTCGGCTGGGGTGG-- 1346
Db	62564 ACCAGTCAAGTGTGACATGCTTGTGTTTATACACAGCTCGGCTGGGGTGGGG 62623
QY	1347 -TTGGAGGCTTTTAAAGGAAGTTACTGTTATAGAGGCTCTAAGATTCACTCATTT 1405
Db	62624 TGGAGAGGCTTTTAAAGGAAGTTACTGTTATAGAGGCTCTAAGATTCACTCATTT 62683
QY	1406 ATTGGCATCTGTTTAAAGTAGATTAGATC 1435
Db	62684 ATTGGCATCTGTTTAAAGTAGATTAGATC 62713

RESULT 8  
AC098613  
LOCUS

AC098613 185437 bp DNA linear PRI 01-AUG-2002

**DEFINITION** Homo sapiens chromosome 3 clone RP11-24F11, complete sequence.  
**ACCESSION** AC098613  
**VERSION** AC098613.2 GI:22038607  
**KEYWORDS** HTG.  
**SOURCE** human.  
**ORGANISM** Homo sapiens  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**REFERENCE** 1 (bases 1 to 185437)  
**AUTHORS** Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z., Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C. and Haugen, E.D.  
**TITLE** Direct Submission  
**JOURNAL** Unpublished  
**REFERENCE** 2 (bases 1 to 185437)  
**AUTHORS** Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (26-OCT-2001) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA  
**REFERENCE** 3 (bases 1 to 185437)  
**AUTHORS** Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z., Saenphimmachak, C., Phelps, K.A., Buckley, D., Kibukawa, M., Raymond, C. and Haugen, E.D.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (01-AUG-2002) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA  
**COMMENT** On Aug 1, 2002 this sequence version replaced gi:16445164.  
 -----  
 Center: University of Washington Genome Center  
 Center Code: UWGC  
 Web site: <http://www.genome.washington.edu>  
 Contact: [uwgchis@u.washington.edu](mailto:uwgchis@u.washington.edu)  
 -----  
 Project Information  
 Center project name: chr-3  
 Center clone name: RP11-24F11 (bc0137)  
 -----  
 Summary Statistics  
 Sequencing vector: plasmid; 100% of reads  
 Chemistry: Dye-terminator ET; 93% of reads  
 Chemistry: Dye-terminator Big Dye; 7% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 184860 bases at least Q40  
 Consensus quality: 185398 bases at least Q30  
 Consensus quality: 185435 bases at least Q20  
 Insert size: 185437; sum-of-contigs  
 Quality coverage: 7.6x in Q20 bases; sum-of-contigs  
 -----  
**Overlapping Sequences:**  
 5': BAC-110P12 U95626, 111014-bp overlap  
 3': RP11-509121 (UWGC:bc0454) AC104304, 61294-bp overlap  
 -----  
**Sequence Quality Assessment:**  
 This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp.  
 Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.  
 -----  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.  
 -----  
**Sequence Validation:**  
 This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and

vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

	BglII	EcoRI	HindIII
SeqDerMap	FgrrPrint	SeqDerMap	FgrrPrint
-----	-----	-----	-----
11786	12040	8837	9423
-----	-----	-----	-----
2067	2065	6	6501
-----	-----	-----	-----
5681	5720	1846	512
-----	-----	-----	-----
3716	3953	4052	449
-----	-----	-----	-----
875	897	1159	6501
-----	-----	-----	-----
910	897	54	10229
-----	-----	-----	-----
2169	2215	560	1025
-----	-----	-----	-----
1705	1653	2287	3176
-----	-----	-----	-----
5763	5720	4905	1054
-----	-----	-----	-----
5844	5720	3049	3998
-----	-----	-----	-----
5061	5001	1377	1100
-----	-----	-----	-----
2625	2640	9903	124
-----	-----	-----	-----
725	<800	1022	2671
-----	-----	-----	-----
1173	1161	12606	1948
-----	-----	-----	-----
5747	5720	866	83
-----	-----	-----	-----
9995	9684	9817	7455
-----	-----	-----	-----
999	999	3598	1305
-----	-----	-----	-----
4541	4503	452	1047
-----	-----	-----	-----
2950	2981	7549	3299
-----	-----	-----	-----
406	<800	2063	5279
-----	-----	-----	-----
3375	3602	5837	6815
-----	-----	-----	-----
13024	13045	13685	435
-----	-----	-----	-----
416	<800	4104	2509
-----	-----	-----	-----
10298	10140	1943	2218
-----	-----	-----	-----
1478	1452	3964	3797
-----	-----	-----	-----
287	<800	2758	7758
-----	-----	-----	-----
205	<800	167	16192
-----	-----	-----	-----
9428	9684	1307	11020
-----	-----	-----	-----
3809	3953	13306	1978
-----	-----	-----	-----
1275	1234	2423	8741
-----	-----	-----	-----
3351	3381	2243	4795





[illegible]

LPGLIFRSQKEGLHYTCSSHPYSOYQFWKNEQTLKIVILGLVPLLVVVCYSGL  
 KILLCRNEKKRRHRAVRLFTIMIVFLWAPYNIIVILNTFOEFFGLNCSNRLL  
 OAMVETLGMTHCCINPLIYAVGKFRNLYLVFQKHAKRCKCCSIFQOEPER  
 ASSVTRSGDELSVGL\*

BASE COUNT 347 a 338 c 297 g 394 t  
 ORIGIN

Query Match 90.3%; Score 1302; DB 9; Length 1376;

Best Local Similarity 97.7%; Pred. No. 0;

Matches 1344; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 GAATGCCCCACAGACCAAGCTCTCCATCTAGTGACAGGAGCAAGCTAGCAGCAAACT 60  
 DB 1 GAATGCCCCACAGACCAAGCTCTCCATCTAGTGACAGGAGCAAGCTAGCAGCAAACT 60  
 QY 61 TCCCTTCACACAAACTTCATTGCTGGCCAAAAGAGAGTAAATCAATGTAGACATC 120  
 DB 61 TCCCTTCACACAAACTTCATTGCTGGCCAAAAGAGAGTAAATCAATGTAGACATC 120  
 QY 121 TATGTAGGCAATTAACCACTTATGATGTATAAACAGTTTGCAATTCATGAGGCAACT 180  
 DB 121 TATGTAGGCAATTAACCACTTATGATGTATAAACAGTTTGCAATTCATGAGGCAACT 180  
 QY 181 AAATACATCTAGGACTTTATAAGATCACTTTTATTTATCCACAGGTTGACACAGA 240  
 DB 181 AAATACATCTAGGACTTTATAAGATCACTTTTATTTATCCACAGGTTGACACAGA 240  
 QY 241 TGGATTATCAAGTGTCAAGTCCATCTATGACATCAATTTATATCATCGGAGCCCTGCC 300  
 DB 241 TGGATTATCAAGTGTCAAGTCCATCTATGACATCAATTTATATCATCGGAGCCCTGCC 300  
 QY 301 AAAAATCAATGTGAAGCAAAATCGCAGCCCGCTCGCTCGCTCTACTACTCTGCTGT 360  
 DB 301 AAAAATCAATGTGAAGCAAAATCGCAGCCCGCTCGCTCGCTCTACTACTCTGCTGT 360  
 QY 361 TCATCTTTGGTTTGTGGCAACATGCTGTCATCTCACTCACTGATTAACCTGCAAAAGC 420  
 DB 361 TCATCTTTGGTTTGTGGCAACATGCTGTCATCTCACTCACTGATTAACCTGCAAAAGC 420  
 QY 421 TGAAGACATGATGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCTTC 480  
 DB 421 TGAAGACATGATGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCTTC 480  
 QY 481 TTACTGTCCCTCTGCGCTCCTATGCTGCGCCGACCTGGACTTTGGAATCAATGT 540  
 DB 481 TTACTGTCCCTCTGCGCTCCTATGCTGCGCCGACCTGGACTTTGGAATCAATGT 540  
 QY 541 GTCAACTCTTTGACAGGCTCTATTTTATAGGCTTCTCTGGAATCTTCTTCATCATCC 600  
 DB 541 GTCAACTCTTTGACAGGCTCTATTTTATAGGCTTCTCTGGAATCTTCTTCATCATCC 600  
 QY 601 TCCGTGACATPCATAGTACCTGGCTGCTGCTCAATGCTGCTGCTTTTAAAGCCAGGA 660  
 DB 601 TCCGTGACATPCATAGTACCTGGCTGCTGCTCAATGCTGCTGCTTTTAAAGCCAGGA 660  
 QY 661 CGGTCACTTTGGGTGTGACAGTGTATCACTGCTGGTGTGCTGCTGCTGCTGCTC 720  
 DB 661 CGGTCACTTTGGGTGTGACAGTGTATCACTGCTGGTGTGCTGCTGCTGCTGCTC 720  
 QY 721 TCCAGGAATCATCTTTACCAATCTCAAAAGAGTCTTCAATACACTGCAGCTCTC 780  
 DB 721 TCCAGGAATCATCTTTACCAATCTCAAAAGAGTCTTCAATACACTGCAGCTCTC 780  
 QY 781 ATTTTCCAT-----ACATTAAGATAGTCACT 808  
 DB 781 ATTTTCCATAGTCAGTATCAATTTCTGGAAGATTTCCAGACATTAAGATAGTCACT 840  
 QY 809 TGGGCTGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 868  
 DB 841 TGGGCTGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 QY 869 TCGTTCGGTGTGAAATGAGAGAGGACAGGCTGTGAGGCTTATCTTCCACATCA 928  
 DB 928 TCGTTCGGTGTGAAATGAGAGAGGACAGGCTGTGAGGCTTATCTTCCACATCA 928

DB 901 TCGTTCGGTGTGAAATGAGAGAGGACAGGCTGTGAGGCTTATCTTCCACATCA 960  
 QY 929 TGAATGTTTATTTCTTCTTGGGCTCCCTACAAATGCTTCTTCTTCCGACACCTTCC 988  
 DB 961 TGAATGTTTATTTCTTCTTGGGCTCCCTACAAATGCTTCTTCTTCCGACACCTTCC 1020  
 QY 989 AGGAATCTTTGGCTGGAATAATTCAGTAGTCTTAACAGGTTGGACCAAGCTATGACAG 1048  
 DB 1021 AGGAATCTTTGGCTGGAATAATTCAGTAGTCTTAACAGGTTGGACCAAGCTATGACAG 1080  
 QY 1049 TGACAGAGACTCTTGGATGACGCACTGCTGCAATCAACCCCATCATCTATGCTTTGTCG 1108  
 DB 1081 TGACAGAGACTCTTGGATGACGCACTGCTGCAATCAACCCCATCATCTATGCTTTGTCG 1140  
 QY 1109 GGGAGAAGTTCAGAACTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1168  
 DB 1141 GGGAGAAGTTCAGAACTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1200  
 QY 1169 GCAATGCTGTTCTATTTTCCAGCAAGGCTCCGAGCGAGCAAGCTCAGTTTACACCC 1228  
 DB 1201 GCAATGCTGTTCTATTTTCCAGCAAGGCTCCGAGCGAGCAAGCTCAGTTTACACCC 1260  
 QY 1229 GATCCACTGGGAGCAGGAATATCTGTGGCTGTGTGACGCACTCAAGTGGGCTGGTG 1288  
 DB 1261 GATCCACTGGGAGCAGGAATATCTGTGGCTGTGTGACGCACTCAAGTGGGCTGGTG 1320  
 QY 1289 ACCAGTCAAGTGTGCACTGCTTAGTTTCTATACACAGCTGGGCTGGGGT 1344  
 DB 1321 ACCAGTCAAGTGTGCACTGCTTAGTTTCTATACACAGCTGGGCTGGGGT 1376

## RESULT 12

AX297794

LOCUS

AX297794

DEFINITION

Sequence 1 from Patent EP1148126.

ACCESSION

AX297794

VERSION

AX297794.1

KEYWORDS

human.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

AUTHORS

Li, X. and Ruben, S.M.

TITLE

Human g-protein chemokine receptor hdnrl0 (crr5 receptor)

JOURNAL

Patent: EP 1148126-A 1 24-OCT-2001;

HUMAN GENOME SCIENCES, INC. (US)

FEATURES

Location/Qualifiers

1..1414

/organism="Homo sapiens"

/db\_xref="taxon:9606"

259..1317

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAD12330.1"

/translation="MDYQVSSPIYDINVTSEPCQKINVKQIAARLPPLYSVLVTFG

FVGNMVLILINCKRLKMTDYLNLAIISDLFLTFVFWAAQADNFGNMQ

LLTGLYFGFSGIFFIILLIDRYLVAVHVAFAKARTVTFGVTISVITWVAVAS

LPGITFRSOKBGLHYTSSHPYSOYQFWKNEQTLKIVILGLVPLLVVVCYSGL

KTLIRCNKKRHRVRLFTIMIVFLWAPYNIIVILNTFOEFFGLNCSNRLL

OAMVETLGMTHCCINPLIYAVGKFRNLYLVFQKHAKRCKCCSIFQOEPER

ASSVTRSGDELSVGL"

BASE COUNT 353 a 342 c 312 g 407 t

ORIGIN

Query Match 89.7%; Score 1293.4; DB 6; Length 1414;

Best Local Similarity 96.8%; Pred. No. 0;

Matches 1345; Conservative 0; Mismatches 16; Indels 32; Gaps 1;

QY 2 AATTCCTCCACAGACCAAGCTCTCCATCTAGTGACAGGAGCAAGCTAGCAGCAAACT 61

DB 21 AATTCCTCCACAGACCAAGCTCTCCATCTAGTGACAGGAGCAAGCTAGCAGCAAACT 80

```
QY 62 CCCTTCACACAAAACCTTCATCTGCTGGCCAAAAGAGAGTTAAATCAATGATGACATCT 121
Db 81 CCCCTCCTACAGAAACCTTCATCTGCTGGCCAAAAGAGAGTTAAATCAATGATGACATCT 140
QY 122 ATGTAGGCAATTAACACCTATGATGATTAACACAGTTTGCAATTCATGAGGAGCAACTA 181
Db 141 ATGTAGGCAATTAACACCTATGATGATTAACACAGTTTGCAATTCATGAGGAGCAACTA 200
QY 182 AATACATCTTAGGACTTTTAAAGAGATCACTTTTATATGACAGAGGTTGGAACAAGAT 241
Db 201 AATACATCTTAGGACTTTTAAAGAGATCACTTTTATATGACAGAGGTTGGAACAAGAT 260
QY 242 GGAATATCAAGTGTCAAGTCAATCATGACATCAATATATATACATCGGAGCCCTGCCA 301
Db 261 GGAATATCAAGTGTCAAGTCAATCATGACATCAATATATATACATCGGAGCCCTGCCA 320
QY 302 AAAAATCAATGTGAAGCAAAATCGACGCGGCTCTCGCTCGGCTCTACTCAGTGGTGT 361
Db 321 AAAAATCAATGTGAAGCAAAATCGACGCGGCTCTCGCTCGGCTCTACTCAGTGGTGT 380
QY 362 CATCTTTGGTTTGTGGCAACATGCTGGTCACTCTCATCTCATGATAAACTGCAAAAGGCT 421
Db 381 CATCTTTGGTTTGTGGCAACATGCTGGTCACTCTCATCTCATGATAAACTGCAAAAGGCT 440
QY 422 GAAGACATGACATGACATCTACCTGCTCAACCTGCCATCTCTGACCTGTTTTTCCTCT 481
Db 441 GAAGACATGACATGACATCTACCTGCTCAACCTGCCATCTCTGACCTGTTTTTCCTCT 500
QY 482 TACTGTCCCTCTGCGGCTCACTATGCTGCCGCCAGTGGGACTTTGGAATACAAATGTG 541
Db 501 TACTGTCCCTCTGCGGCTCACTATGCTGCCGCCAGTGGGACTTTGGAATACAAATGTG 560
QY 542 TCACACTCTGACAGGCTCTATTTTATAGGCTTCTCTCTGGAATCTCTTCATCAPCCT 601
Db 561 TCACACTCTGACAGGCTCTATTTTATAGGCTTCTCTCTGGAATCTCTTCATCAPCCT 620
QY 602 CCTGCAATCGATAGTACTGCTGCTGCTCCATGCTGTTGCTTTAAAGCCAGGAC 661
Db 621 CCTGCAATCGATAGTACTGCTGCTGCTCCATGCTGTTGCTTTAAAGCCAGGAC 680
QY 662 GGTCACTTTGGGTTGTGACAGTGTGATCACTTGGGTTGGTGGTGTGTTGGTCTCT 721
Db 681 GGTCACTTTGGGTTGTGACAGTGTGATCACTTGGGTTGGTGGTGTGTTGGTCTCT 740
QY 722 CCCAGGAATCATCTTTACAGATCTCAAAAAGAGGTTCTCAATACACCTGCAGCTCTCA 781
Db 741 CCCAGGAATCATCTTTACAGATCTCAAAAAGAGGTTCTCAATACACCTGCAGCTCTCA 800
QY 782 TTTTCCAT-----ACATTAAGATAGTCACTT 809
Db 801 TTTTCCATAGTCACTTCAATTCGGAAGAAATTCACAGATTAAGATAGTCACTT 860
QY 810 GGGGCTGGTCTCGCGCTGCTTGTGATGCTATGCTGCTACTCGGGAATCTCAAAACTCT 869
Db 861 GGGGCTGGTCTCGCGCTGCTTGTGATGCTATGCTGCTACTCGGGAATCTCAAAACTCT 920
QY 870 GCTTCGGTGTCAAAATGAGAGAGAGGACAGGCGCTGTGAGGCTTATCTTCACCATCAT 929
Db 921 GCTTCGGTGTCAAAATGAGAGAGAGGACAGGCGCTGTGAGGCTTATCTTCACCATCAT 980
QY 930 GATTGTTTATTTCTTCTGCGCTCCCTFACAAATGCTCTCTCTCGAACACCTTCCA 989
Db 981 GATTGTTTATTTCTTCTGCGCTCCCTFACAAATGCTCTCTCTCGAACACCTTCCA 1040
QY 990 GGAATCTTTGGCCTGAATATTCAGTACTCTAACAGGTTGGACCAAGCTATGACAGT 1049
Db 1041 GGAATCTTTGGCCTGAATATTCAGTACTCTAACAGGTTGGACCAAGCTATGACAGT 1100
QY 1050 GACAGACTCTTTGGGATGAGGACCTGCTGATCAACCCCATCATCTATGCTTTGCGG 1109
Db 1101 GACAGACTCTTTGGGATGAGGACCTGCTGATCAACCCCATCATCTATGCTTTGCGG 1160
```

```
1110 GGAGAAGTTCAGAACTACCTCTTAGTCTTCTTCCAAAGCACATTCGCAAAAGCTTCTG 1169
1161 GGAGAAGTTCAGAACTACCTCTTAGTCTTCTTCCAAAGCACATTCGCAAAAGCTTCTG 1220
1170 CAATGCTGTTCTATTTTCCAGCAAGAGGCTCCGAGCGGCAAGCTCAGTTTACACCG 1229
1221 CAATGCTGTTCTATTTTCCAGCAAGAGGCTCCGAGCGGCAAGCTCAGTTTACACCG 1280
1230 ATCCACTGGGAGCAGCAAAATATCTGTGGCTTGTGACACGGACTCAAGTGGCTGGTGA 1289
1281 ATCCACTGGGAGCAGCAAAATATCTGTGGCTTGTGACACGGACTCAAGTGGCTGGTGA 1340
1290 CCCAGTCAAGTGTGTCACATGCTTAGTCTTTCATACACACCTCGGCTGGGCTGGTGG 1349
1341 CCCAGTCAAGTGTGTCACATGCTTAGTCTTTCATACACACCTCGGCTGGGCTGGGCT 1400
1350 GGAGGCTCTTTTTT 1362
1401 GGAAGAGTCTTTT 1413
```

RESULT 13  
AX297803  
LOCUS AX297803 1414 bp DNA linear PAT 22-NOV-2001  
DEFINITION Sequence 1 from Patent EP1149582.  
ACCESSION AX297803  
VERSION AX297803.1 GI:17065853  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Li, Y. and Ruben, S.M.  
Human q-protein chemokine receptor hdnr10 (ccr5 receptor). Uses  
thereof  
Patent: EP 1149582-A 1 31-OCT-2001;  
HUMAN GENOME SCIENCES, INC. (US)  
LOCATION/Qualifiers  
1. 1414  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
259..1317  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD12331.1"  
/db\_xref="GI:17065853"  
/translation="MDYQVSPYIDNYITSEPCQKINKVQIAARLLPPLYSLVFIFG  
FVGNMLVILLINCKRLKSMVDIYLLNLALSDLEFLITVPWHAHYAAQWDFGNTKQ  
LLGLFIFGFSGFIFILLIDRYLVAVHAFKARTVTVGVISVITVWVAFAS  
LPLIFTRSKGGLHYTSSHPYQYQWKNFQTLKIVILGLVLPPLVWVYSGIL  
KTLRCNRKKHRAVRIFTIMIVYFLWAPYINIVLLNTQEFGLNCSNSRLD  
QAMQVTELTGMPHCCINPIIYAFVGEKFNVLVFFOKHAKRFCKGCSIFQDPAPER  
ASSVYTRSTSEQIEISVGL"  
BASE COUNT 353 a 342 c 312 g 407 t  
ORIGIN

Query Match 89.7%; Score 1293.4; DB 6; Length 1414;  
Best Local Similarity 96.6%; Pred. No. 0;  
Matches 1345; Conservative 0; Mismatches 16; Indels 32; Gaps 1;

QY 2 AATCCCCCAGGAGCCAGCTCTCCATCTAGTGACAGGAGGAGTAAATCAATGATGACATCT 61  
Db 21 ATCTCCCAAGAGCCAGCTCTCCATCTAGTGACAGGAGGAGTAAATCAATGATGACATCT 80  
QY 62 CCCTTCACACAAAACCTTCATCTGCTGGCCAAAAGAGAGTTAAATCAATGATGACATCT 121  
Db 81 CCCTTCACAGAACTTCATCTGCTGGCCAAAAGAGAGTTAAATCAATGATGACATCT 140  
QY 122 ATGTAGGCAATTAACACCTATGATGATTAACACAGTTTGCAATTCATGAGGAGCAACTA 181  
Db 141 ATGTAGGCAATTAACACCTATGATGATTAACACAGTTTGCAATTCATGAGGAGCAACTA 200

Db	1281	ATCCACTGAGGAGCGAAATATCTGTGGGCTTGTCACAGGACTCAAGTGGGCTGGTGA	1340
Qy	1290	CCCAAGTCAGAGTGTGCACATGGCTTAGTTTTCATACACAGCCTGGGCTGGGGTGGTGG	1349
Db	1341	CCCAAGTCAGAGTGTGCACATGGCTTAGTTTTCATACACAGCCTGGGCTGGGGTGGGGT	1400
Qy	1350	GGAGGCTCTTTTTT	1362
Db	1401	GGAGAGGTCTTT	1413
RESULT 14	AX297812		
LOCUS	AX297812	1414 bp	DNA linear PAT 22-NOV-2001
DEFINITION	Sequence 1 from Patent EP1146122.		
ACCESSION	AX297812		
VERSION	AX297812.1	GI:17065855	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE 1	Li, Y. and Ruben, S.M.		
AUTHORS	Human g-protein chemokine receptor hGnrl10 (ccr5 receptor)		
TITLE	Patent: EP 1146122-A 1 17-OCT-2001;		
JOURNAL	HUMAN GENOME SCIENCES, INC. (US)		
FEATURES	location/Qualifiers		
source	1..1414		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
CDS	259..1317		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAD12332.1"		
	/translation="MDYQVSPIYDINTYSEPCQKINYKQIAARLLPPLYSLVIFGIFV		
	FGVNEVILLLNCKRLKMSDTIILNLAISDLFLLTPVFWAHYAAAGQWDEGNTMQ		
	LLTGIFGIFGFFIILLTIDRLAVYHVAFLAKARTVTFGVVTSVITWVAFAAS		
	LPGLIFTRCKGGLHYTCSSHPYQYOFKNFTOLKIVILGLVLPILVWVICYGL		
	KTLRLCNKRRHRAVLIETIMIVYFLPNAPINIVLLINTPOEFGNLCNCSNNRLD		
	QAMQYETLQNTCHCLINPIIYAEVGEKFRNVLVFFQKHIAKRCCKCSIFQGEAPER		
	ASSYITNSDEEIGSVGL"		
BASE COUNT	353 a	342 c	312 g
ORIGIN	407 t		
Query Match	89.7%;	Score 1293.4;	DB 6; Length 1414;
Best Local Similarity	96.6%;	Pred. No. 0;	
Matches 1345;	Conservative	0; Mismatches	16; Indels 32; Gaps 1;
Qy	2	AATTCOCCCAACAGACGCCAAGCTCCATCTAGTGGACAGGGAAGCTAGCAGCAAAACCTT	61
Db	21	AATTCOCCCAACAGACGCCAAGCTCCATCTAGTGGACAGGGAAGCTAGCAGCAAAACCTT	80
Qy	62	CCCTTCACACAAACCTTCATGCTTGGCCAAAAGAGAGTTAATTCAGTGAACATCT	121
Db	81	CCCTTCACACAAACCTTCATGCTTGGCCAAAAGAGAGTTAATTCAGTGAACATCT	140
Qy	122	ATGTAGGCAATTAACAACTTATGATGATATAAAACAGTTTGATTCATCGAGGAGCAACTA	181
Db	141	ATGTAGGCAATTAACAACTTATGATGATATAAAACAGTTTGATTCATCGAGGAGCAACTA	200
Qy	182	AATACATTTCTAGGACTTTATAAAAGATCACTTTTATTTATGACAGGGTGGAAACAAGT	241
Db	201	AATACATTTCTAGGACTTTATAAAAGATCACTTTTATTTATGACAGGGTGGAAACAAGT	260
Qy	242	GGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATATATATACATCGGAGCCCTGGCA	301
Db	261	GGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATATATATACATCGGAGCCCTGGCA	320
Qy	302	AAAAATCAATGTGAAGCAAAATCGGACCGCGCCTCTCGCTCCGCTCTACTCACTGGGT	361



Db 321 AAAATCAATGTGAAGCAAAATCGAGCCCGCTCTCGCTCGCTCTACTCATCTGGTGT 380  
QY 362 CATCTTTGGTTTTTGTGGCAACATGCTGGTCATCCATCCTGATGATAAACTGCAAAAGGCT 421  
Db 381 CATCTTTGGTTTTTGTGGCAACATGCTGGTCATCCCTGATGATAAACTGCAAAAGGCT 440  
QY 422 GAAGACATGACTGACATGACTGCTCAACCTGGCGACTCTGACCTGTTTTTCTCTCT 481  
Db 441 GAAGACATGACTGACATGACTGCTCAACCTGGCGACTCTGACCTGTTTTTCTCTCT 500  
QY 482 TACTGCCCTCTTGGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 541  
Db 501 TACTGCCCTCTTGGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 560  
QY 542 TCAACTCTTGACAGGCTCTATTTATAGGCTTCTCTGGAATCTTCTCATCATCT 601  
Db 561 TCAACTCTTGACAGGCTCTATTTATAGGCTTCTCTGGAATCTTCTCATCATCT 620  
QY 602 CTTGACAAATGATAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661  
Db 621 CTTGACAAATGATAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 680  
QY 662 GGTACCTTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 721  
Db 681 GGTACCTTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 740  
QY 722 CCAGAGATCATCTTACAGATCTCAAAAGAGGCTTCTCATACCTGCACTCTCA 781  
Db 741 CCAGAGATCATCTTACAGATCTCAAAAGAGGCTTCTCATACCTGCACTCTCA 800  
QY 782 TTTTCCAT-----ACATTAAGATATCATCT 809  
Db 801 TTTTCCATACAGTACATCAATCTCGAAGAAATTTCCAGACATTAAGATATGCTATCT 860  
QY 810 GGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 869  
Db 861 GGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920  
QY 870 GCTTGGGTGCGAATGAGAAGAGGACAGGCTGAGGCTTATCTTCCACCATCAT 929  
Db 921 GCTTGGGTGCGAATGAGAAGAGGACAGGCTGAGGCTTATCTTCCACCATCAT 980  
QY 930 GATGTTTATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 989  
Db 981 GATGTTTATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1040  
QY 990 GGAATCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1049  
Db 1041 GGAATCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1100  
QY 1050 GACAGAGACTCTGGGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1109  
Db 1101 GACAGAGACTCTGGGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1160  
QY 1110 GGAGAGTTGAGAACTTACCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1169  
Db 1161 GGAGAGTTGAGAACTTACCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1220  
QY 1170 CAATGCTGTTCTATTTTCCAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1229  
Db 1221 CAATGCTGTTCTATTTTCCAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1280  
QY 1230 ATCCATCTGGGAGCAGAAATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1289  
Db 1281 ATCCATCTGGGAGCAGAAATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1340  
QY 1290 CCCAGTCAAGAGTTGTCACATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1349  
Db 1341 CCCAGTCAAGAGTTGTCACATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1400  
QY 1350 GGAGGCTTTTTT 1362  
Db 1401 GGAAGAGGCTTTT 1413

RESULT 15  
AX297821  
LOCUS 1414 bp DNA linear PAT 22-NOV-2001  
DEFINITION Sequence 1 from Patent EP1145721.  
ACCESSION AX297821  
VERSION AX297821.1 GI:17065857  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Li, Y. and Ruben, S. M.  
Human g-protein chemokine receptor hGnrl10 (ccr5 receptor).  
Pharmaceutical composition  
Patent: EP 1145721-A 1 17-OCT-2001;  
HUMAN GENOME SCIENCES, INC. (US)  
FEATURES  
Location/Qualifiers  
1..1414  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
259..1317  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="GI:17065858"  
/translation="MDYQSSPIYDINVTSEPCQKINVKQIAARILPLYSIVPIFG  
FVGNMLVILILNCKRLKSWIDILNLAIISDLFLVTPFWAHYAAQWDFGNTWCQ  
LUTGYFFIGSGIFILLTIDRLAVHAFKARTVTVGVTVSVITWVAVFAS  
LGIIFTRSKQEGHLVTCSSHPYQYQFWKQFLKILVGLVPLLVNVCSSGIL  
KELLRCRKKRHRVLLFTIMVIFLWAPYINVLINTFOEFLNCCSSNRDL  
QAMVTEFLGTHCCINPIIYAVGKFRNLLVFFQXHIKRFCKCISFQQAEPER  
OASVTRSTEEQETSVGL"  
BASE COUNT 353 a 342 c 312 g 407 t  
ORIGIN  
Query Match 89.7%; Score 1293.4; DB 6; Length 1414;  
Best Local Similarity 96.6%; Pred. No. 0;  
Matches 1345; Conservative 0; Mismatches 16; Indels 32; Gaps 1;  
QY 2 AATTCCTCCCAAGAGAGGCTCTCCATCTAGTGGAGAGGAGCTAGCAGCAAACTT 61  
Db 21 ATTCCCTCCCAAGAGAGGCTCTCCATCTAGTGGAGAGGAGCTAGCAGCAAACTT 80  
QY 62 CCCTTCCTACTACAAACTTCTGCTGGCCAAAAGAGAGTTAATTCATGTAGACATCT 121  
Db 81 CCCTTCCTACTACAAACTTCTGCTGGCCAAAAGAGAGTTAATTCATGTAGACATCT 140  
QY 122 ATGTAGCAATTAATAACCTTATGATGATATAAAGCTTTGCATTCATGGAGGCAACTA 181  
Db 141 ATGTAGCAATTAATAACCTTATGATGATATAAAGCTTTGCATTCATGGAGGCAACTA 200  
QY 182 AATACATCTTAGGCTTTATAAAGATCACTTTTATTTATGACAGGCTGGAACAGAT 241  
Db 201 AATACATCTTAGGCTTTATAAAGATCACTTTTATTTATGACAGGCTGGAACAGAT 260  
QY 242 GGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATTTATATACATGGAGCCTGCCA 301  
Db 261 GGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATTTATATACATGGAGCCTGCCA 320  
QY 302 AAAATCAATGTGAAGCAAAATCGAGCCCGCTCTCGCTCGCTCTACTCATCTGGTGT 361  
Db 321 AAAATCAATGTGAAGCAAAATCGAGCCCGCTCTCGCTCGCTCTACTCATCTGGTGT 380  
QY 362 CATCTTTGGTTTTTGTGGCAACATGCTGGTCATCCCTGATGATAAACTGCAAAAGGCT 421  
Db 381 CATCTTTGGTTTTTGTGGCAACATGCTGGTCATCCCTGATGATAAACTGCAAAAGGCT 440  
QY 422 GAAGACATGACTGACATGACTGCTCAACCTGGCGACTCTGACCTGTTTTTCTCTCT 481  
Db 441 GAAGACATGACTGACATGACTGCTCAACCTGGCGACTCTGACCTGTTTTTCTCTCT 500



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 00:19:22 ; Search time 357 Seconds  
(without alignments)  
9096.307 Million cell updates/sec

Title: US-09-938-703-3

Perfect score: 1442

Sequence: 1 GAATCCCCCAGAGAGCA.....AGTAGATGATCGAATTC 1442

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	1442	100.0	1442	18	AAT90118	CDNA for inactive
2	1382.8	95.9	1477	18	AAT90117	CDNA for human CCR
3	1382.8	95.9	1477	22	AAT87099	Human CCR5 cDNA se
4	1375.2	95.4	9141	24	ABA97318	Human chemokine (C
5	1375.2	95.4	143068	21	AAT21105	Human low adenosin
6	1375.2	95.4	143068	21	AAT21272	Human low adenosin
7	1375.2	95.4	143068	21	AAA35150	Human adenosine re
8	1375.2	95.4	143068	21	ABA68124	Ovary cancer relat
9	1375.2	95.4	149412	21	AAA35151	Human adenosine re

10	1375.2	95.4	152740	21	AAF21273	Human low adenosin
11	1373.6	95.3	143068	21	AAA34983	Human adenosine re
12	1344	93.2	1344	20	AAV84159	HIV-1 co-receptor
13	1302	90.3	1376	22	AAH26903	Human HIV-1 co-rec
14	1299	90.1	1376	22	AAV84126	HIV-1 co-receptor
15	1293.4	89.7	1414	22	AAF26390	Human HDGNR10 cDNA
16	1288.6	89.4	1414	18	AAT44042	Human G-protein ch
17	1288.6	89.4	1414	21	AAZ91481	Human G-protein ch
18	1288.6	89.4	1414	22	AAI3181	Human G-protein ch
19	1288.6	89.4	1414	22	AAI3181	Human G-protein ch
20	1288.6	89.4	1414	24	AAI3181	Human G-protein ch
21	1160	80.4	1557	18	AAI3181	DNA encoding human
22	1149	79.7	3383	18	AAI3181	Human chemokine re
23	1149	79.7	3383	21	AAI3181	Human chemokine re
24	1149	79.7	3383	21	AAI3181	Human low adenosin
25	1149	79.7	3383	21	AAI3181	Human adenosine re
26	1117	77.5	1225	19	AAI3181	Human chemokine re
27	1117	77.5	1225	19	AAI3181	DNA encoding human
28	1115.4	77.4	1225	24	AAI3181	Human CC chemokine
29	1100	76.3	1255	19	AAI3181	Human CCR5 Gln 55
30	985	68.3	1059	19	AAI3181	DNA encoding human
31	985	68.3	1059	24	AAI3181	Human CC-CCR5 codi
32	983.4	68.2	1071	20	AAI3181	Human chemokine (C
33	982.8	68.2	5674	20	AAI3181	HIV-1 co-receptor
34	981.8	68.1	1059	23	AAI3181	Human chemokine re
35	980.4	68.0	1056	22	AAI3181	Non-endogenous hum
36	980.4	68.0	1056	22	AAI3181	Human G-protein ch
37	980.4	68.0	1056	24	AAI3181	Human G-protein ch
38	954.6	66.2	1059	22	AAI3181	DNA encoding human
39	948.2	65.8	1059	18	AAI3181	Nucleotide sequenc
40	792	54.9	792	18	AAI3181	Macaque chemokine
41	727.4	50.4	2440	19	AAI3181	CDNA for inactive
42	683.4	47.4	1979	16	AAI3181	Mouse CC-CCR5 codi
43	681.6	47.3	1083	22	AAI3181	Human monocyte che
44	680	47.2	1083	22	AAI3181	Human CCR2-641 pol
45	676.8	46.9	1083	18	AAI3181	Human wild-type CC
						Human monocyte che

# ALIGNMENTS

RESULT 1  
AAT90118  
ID AAT90118 standard; cDNA; 1442 BP.

XX AAT90118;

XX AC AAT90118;

XX DT 14-APR-1998 (first entry)

XX DE CDNA for inactive human CCR5.

XX KW Inactive; human Cys-Cys chemokine receptor 5; CCR5;  
KW human immunodeficiency virus; type 1; type 2; HIV-1; HIV-2;  
KW predisposition; resistance; diagnosis; treatment; prevention;  
KW inflammatory disease; rheumatoid arthritis; glomerulonephritis;  
KW asthma; idiopathic pulmonary fibrosis; psoriasis; viral infection;  
KW cancer; atherosclerosis; autoimmune disorder; ss.

XX OS Homo sapiens.

XX FE Key Location/Qualifiers  
XX FT CDS 240..887  
XX FT /\*tag= a

XX PN W09732019-A2.

XX XX

XX PD 04-SEP-1997.

XX XX

XX PF 28-FEB-1997; 97WO-BE00023.

XX XX

XX PR 06-AUG-1996; 96EP-0870102.

XX PR 01-MAR-1996; 96EP-0870021.

XX XX



KW cancer; atherosclerosis; autoimmune disorder; ss.  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FH 240..1298  
FT CDS /\*tag= a  
FT  
XX  
XX W09732019-A2.  
XX  
XX 04-SEP-1997.  
XX 28-FEB-1997; 97WO-BE00023.  
XX 06-AUG-1996; 96EP-0870102.  
XX 01-MAR-1996; 96EP-0870021.  
XX (EURO-) EUROSREEN SA.  
XX  
XX Libert F, Parmentier M, Samson M, Vassart G;  
XX WPI: 1997-479829/44.  
XX P-PSDB; AAT90117.  
XX  
XX Active and inactive forms of human CC chemokine receptor CCR-5 -  
XX useful to diagnose, prevent and/or treat inflammatory disorders,  
XX autoimmune disease and viral infection  
XX  
XX Claim 18; Fig 1b-c; 94pp; English.  
XX  
XX The present sequence encodes human CC (Cys-Cys) chemokine receptor  
XX 5 (CCR5), which is stimulated by MIP-1 alpha, MIP-1 beta or RANTES  
XX chemokines, but not by monocyte chemoattractant protein 1 (MCP-1),  
XX MCP-2, MCP-3, interleukin-8 (IL-8) or growth related gene product  
XX alpha (GRO alpha) chemokines. Active CCR-5 is also a receptor of  
XX human immunodeficiency virus type 1 or type 2 (HIV-1 or HIV-2).  
XX CCR5 or its cDNA can be used to diagnose, treat and/or prevent  
XX inflammatory diseases, e.g. rheumatoid arthritis,  
XX glomerulonephritis, asthma, idiopathic pulmonary fibrosis and  
XX psoriasis, viral infections, especially HIV-1 or HIV-2 infection,  
XX cancer, atherosclerosis and autoimmune disorders.  
XX  
XX Sequence 1477 BP; 374 A; 349 C; 320 G; 431 T; 3 other;  
SQ  
Query Match 95.98; Score 1382.8; DB 18; Length 1477;  
Best Local Similarity 97.48; Pred. No. 0;  
Matches 1439; Conservative 0; Mismatches 3; Indels 35; Gaps 2;  
QY 1 GAATTCCTCCACAGAGCCAGCTCTCCATCTAGTGACAGGAGGAGTAAATCATGTAGACATC 60  
DB 1 GAATTCCTCCACAGAGCCAGCTCTCCATCTAGTGACAGGAGGAGTAAATCATGTAGACATC 60  
QY 61 TCCTTCATCAAACTTCATGTGCTGGCCAAAAGAGAGTAAATCATGTAGACATC 120  
DB 61 TCCTTCATCAAACTTCATGTGCTGGCCAAAAGAGAGTAAATCATGTAGACATC 120  
QY 121 TATGTAGCAATTAATAACCTATTGATGTATATAAAGAGTTTGCATTATGAGGAGCAACT 180  
DB 121 TATGTAGCAATTAATAACCTATTGATGTATATAAAGAGTTTGCATTATGAGGAGCAACT 180  
QY 181 AATATCAATTCAGGACTTTATAAAGATCACTTTTATTTATGACAGGTTGGAACAAGA 240  
DB 181 AATATCAATTCAGGACTTTATAAAGATCACTTTTATTTATGACAGGTTGGAACAAGA 240  
QY 241 TGAATATCAAGTGTCAAGTCCAACTATGACATCAATATTTATATCATCGAGGCCCTGCC 300  
DB 241 TGAATATCAAGTGTCAAGTCCAACTATGACATCAATATTTATATCATCGAGGCCCTGCC 300  
QY 301 AAAAAATCAATGTAGCAAAATCGCAGCCGCTCCCTCGCTCTACTCACTGGTGT 360  
DB 301 AAAAAATCAATGTAGCAAAATCGCAGCCGCTCCCTCGCTCTACTCACTGGTGT 360  
QY 361 TCATCTTTGGTTTGGGCAACATCTGGTCTATCTCTCATCTGATATAAATCGCAAAAGGC 420

DB 361 TCATCTTTGGTTTGGGCAACATCTGGTCTATCTCTCATCTGATAAAATCGAAAAGC 420  
QY 421 TGAAGAGCATGACTGACATCTACCTGCTCAACTGGCCATCTCTGACCTGTTTTCCTTC 480  
DB 421 TGAAGAGCATGACTGACATCTACCTGCTCAACTGGCCATCTCTGACCTGTTTTCCTTC 480  
QY 481 TTACTGTCCCTTCTGGGCTCACTATGCTGCCGCCACGTGGGACTTTGGAATCAATGT 540  
DB 481 TTACTGTCCCTTCTGGGCTCACTATGCTGCCGCCACGTGGGACTTTGGAATCAATGT 540  
QY 541 GTCAACTCTTGACAGGGCTCTATTTATAGGCTTCTCTGGAATCTTCTTCATCATCC 600  
DB 541 GTCAACTCTTGACAGGGCTCTATTTATAGGCTTCTCTGGAATCTTCTTCATCATCC 600  
QY 601 TCCTGACAATCGATAGTACCTGGCTGCTGCTCACTGCTGTTTGTGTTTAAAGCCAGGA 660  
DB 601 TCCTGACAATCGATAGTACCTGGCTGCTGCTCACTGCTGTTTGTGTTTAAAGCCAGGA 660  
QY 661 CGGTCACTTTGGGCTGGTGACAAAGTGTATCACTTTGGGCTGGTGTGGCTTC 720  
DB 661 CGGTCACTTTGGGCTGGTGACAAAGTGTATCACTTTGGGCTGGTGTGGCTTC 720  
QY 721 TCCAGGAATCATCTTTACCAGATCTCAAAAAGAGGCTTTCATTAACCTGCACTCTC 780  
DB 721 TCCAGGAATCATCTTTACCAGATCTCAAAAAGAGGCTTTCATTAACCTGCACTCTC 780  
QY 781 ATTTTCCAT-----ACATTAAGATAGTCACT 808  
DB 781 ATTTTCCATAGCTAGTATCAATCTGGAAGAAATTCACAGACATTAAGATAGTCACT 840  
QY 809 TGGGCTGGTCTGCGCTGCTGCTCATGGTCACTGCTACTCGGAATCTTAAACTC 868  
DB 841 TGGGCTGGTCTGCGCTGCTGCTCATGGTCACTGCTACTCGGAATCTTAAACTC 900  
QY 869 TGTCTTCGGTGTCAAAATGAGAGAGGACAGGGCTGTGAGGCTTATCTTCAACATCA 928  
DB 901 TGTCTTCGGTGTCAAAATGAGAGAGGACAGGGCTGTGAGGCTTATCTTCAACATCA 960  
QY 929 TGATGTTTATTTCTTCTTGGGCTCCCTACACATTTGCTTCTCTGACACCTTC 988  
DB 961 TGATGTTTATTTCTTCTTGGGCTCCCTACACATTTGCTTCTCTGACACCTTC 1020  
QY 989 AGGAATCTTTGGCTGAAATTAATGACAGTAGCTCTACAGGTTGGACCAAGCTATGAGG 1048  
DB 1021 AGGAATCTTTGGCTGAAATTAATGACAGTAGCTCTACAGGTTGGACCAAGCTATGAGG 1080  
QY 1049 TGACAGAGACTCTTGGGATGACGACTGTGCTCAACCCCATCATCTATGCTTCTGCG 1108  
DB 1081 TGACAGAGACTCTTGGGATGACGACTGTGCTCAACCCCATCATCTATGCTTCTGCG 1140  
QY 1109 GGGAGAGTTCAAGAACTACCTCTTAGTCTTCTTCCAAAAGACATTTGCCAAGCTTCT 1168  
DB 1141 GGGAGAGTTCAAGAACTACCTCTTAGTCTTCTTCCAAAAGACATTTGCCAAGCTTCT 1200  
QY 1169 GCAATCTCTCTTATTTTCCAGCAAGAGGCTCCCGAGCGAGCAAGCTCAGTTTACACC 1228  
DB 1201 GCAATCTCTCTTATTTTCCAGCAAGAGGCTCCCGAGCGAGCAAGCTCAGTTTACACC 1260  
QY 1229 GATCCACTGGGAGGAGGAGAAATATCTTGGGCTTGTGACGGACTCAAGTGGGCTGGT 1288  
DB 1261 GATCCACTGGGAGGAGGAGAAATATCTTGGGCTTGTGACGGACTCAAGTGGGCTGGT 1320  
QY 1289 ACCAGTCAAGTGTGACATAGGCTTAGTCTTATACACAGCTGGGCTGGGCTGGT 1348  
DB 1321 ACCAGTCAAGTGTGACATAGGCTTAGTCTTATACACAGCTGGGCTGGGCTGGT 1380  
QY 1349 GG---GAGGCTTTTAAAGAGAGTACTGTTATAGAGGCTTAAGATTCATCCATTT 1405  
DB 1381 TGGNAGGCTTTTAAAGAGAGTACTGTTATAGAGGCTTAAGATTCATCCATTT 1440  
QY 1406 ATTTGGCATCTGTTTAAAGTAGATTAGTCCGAATTC 1442

DB	1441	ATTGGCATCTGTTAAAGTAGATAGATCCGAATTC	1477
RESULT 3			
AAF87099			
ID	AAF87099	standard; cDNA; 1477 BP.	
XX			
AC	AAF87099;		
DT	09-OCT-2001	(first entry)	
XX			
DE	Human CCR5	cDNA sequence.	
XX			
KW	Chemotactic chemokine receptor 5; gp120; CD4; therapy; HIV; CCR5;		
KW	human immunodeficiency virus; anti-inflammatory disease; human; ss.		
XX			
OS	Homo sapiens.		
XX			
PH	Key	Location/Qualifiers	
FT	CDS	240..1298	
FT		/*tag= a	
FT		/product= "CCR5"	
PN	EP1118858-A2.		
XX			
PD	25-JUL-2001.		
XX			
PF	03-JAN-2001; 2001EP-0300020.		
XX			
PR	12-JAN-2000; 2000GB-0000659.		
PR	12-JAN-2000; 2000GB-0000661.		
PR	12-JAN-2000; 2000GB-0000663.		
XX			
PA	(PFIZ ) PFIZER LTD.		
PA	(PFIZ ) PFIZER INC.		
XX			
PI	Dobbs S, Perros M, Rickett GA;		
XX			
DR	WPI; 2001-477088/52.		
DR	P-PSDB; ANB8335A.		
XX			
PT	Determining if an agent can modulate CCR5-gp120 interaction, comprises		
PT	incubating the agent with CCR5 and gp120 and determining if the agent		
PT	modulates the interaction		
XX			
PS	Claim 1; Page 109-110; 113pp; English.		
XX			
CC	This sequence represents the human CCR5 cDNA sequence.		
CC	The invention relates to a method for determining whether an agent is		
CC	capable of modulating the interaction of chemotactic chemokine receptor 5		
CC	(CCR5) with gp120, comprising incubating the agent with CCR5 and gp120		
CC	and determining whether the agent modulates the interaction, where gp120		
CC	is associated with CD4, and where the interaction is a low affinity		
CC	binding. The method is used to identify an agent capable of modulating		
CC	the interaction of CCR5 with gp120. An agent identified by the method		
CC	is used to prepare a pharmaceutical composition for the treatment of a		
CC	disease or condition associated with CCR5 and gp120 interaction, to treat		
CC	a subject with a disease or condition associated with CCR5 and gp120		
CC	interaction, and for preparing a pharmaceutical for treating human		
CC	immunodeficiency virus (HIV). It can also be used to treat anti-inflammatory		
CC	diseases. The method is commercially useful, amenable to high throughput		
CC	screening, and detects interaction of gp120 with cells expressing only		
CC	CCR5.		
XX			
SQ	Sequence 1477 BP; 374 A; 349 C; 320 G; 431 T; 3 other;		
Query Match	95.9%;	Score 1382.8;	DB 22; Length 1477;
Best Local Similarity	97.4%;	Pred. No. 0;	
Matches 1439;	Conservative	0; Mismatches	3; Indels 35; Gaps 2;
QY	1	GAATTCCTCCCAACAGAGCCAGCTCTCCATCTAGTGGACAGGAACTAGCAGCAACCT	60
DB	1	GAATTCCTCCCAACAGAGCCAGCTCTCCATCTAGTGGACAGGAACTAGCAGCAACCT	60

```
QY 1109 GGGAGAGTTCAGAACTACCTCTTAGCTCTTCCAAAAGCACATTGCCAAACGGTTCT 1168
    |||
Db 1141 GGGAGAGTTCAGAACTACCTCTTAGCTCTTCCAAAAGCACATTGCCAAACGGTTCT 1200
    |||
QY 1169 GCAAATGCTGTTCTATTTCCAGCAAGAGGCTCCCGAGCGAGCAAGCTCAGTTTACACCC 1228
    |||
Db 1201 GCAAATGCTGTTCTATTTCCAGCAAGAGGCTCCCGAGCGAGCAAGCTCAGTTTACACCC 1260
    |||
QY 1229 GATCCACTGGGAGCAGGAAATATCTGTGGGCTTGTGACACGAGCTCAAGTGGGCTGGTG 1288
    |||
Db 1261 GATCCACTGGGAGCAGGAAATATCTGTGGGCTTGTGACACGAGCTCAAGTGGGCTGGTG 1320
    |||
QY 1289 ACCAGTCAGAGTGTGCACATGGCTTAGTTTTCATACACAGCCGCGGCTGGGGTGGT 1348
    |||
Db 1321 ACCAGTCAGAGTGTGCACATGGCTTAGTTTTCATACACAGCCGCGGCTGGGGTGGT 1380
    |||
QY 1349 GG---GAGGCTCTTTTAAAGGAAGTACTGTTATAGAGGGTCTAAGATTTCATCCATT 1405
    |||
Db 1381 TGGNAGGCTCTTTTAAAGGAAGTACTGTTATAGAGGGTCTAAGATTTCATCCATT 1440
    |||
QY 1406 ATTGGCATCTGTTTAAAGTAGATATAGATCCGAATTC 1442
    |||
Db 1441 ATTGGCATCTGTTTAAAGTAGATATAGATCCGAATTC 1477
    |||

RESULT 4
ID ABA97318 standard; DNA; 9141 BP.
AC ABA97318;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human chemokine (C-C motif) receptor 5 gene.
XX
KW Human chemokine C-C motif receptor 5; CCR5; haplotype pair; isogene;
KW single nucleotide polymorphism; SNP; human immunodeficiency virus 1;
KW HIV-1 infection; acquired immunodeficiency syndrome; AIDS; antiviral;
KW genotype; polymorphic variant; transgenic; drug screening; gene therapy;
KW gene; chromosome 3p21; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace (1629,A)
FT /*tag= a
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace (1953,C)
FT /*tag= b
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace (1956,T)
FT /*tag= c
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace (2002,A)
FT /*tag= d
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace (2253,T)
FT /*tag= e
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace (2402,C)
FT /*tag= f
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace (3920,T)
FT /*tag= g
FT /*standard_name= "Single nucleotide polymorphism"
FT CDS 4083..5141
FT /*tag= h
FT /*product= "CCR5"
FT /*note= "Human chemokine (C-C motif) receptor 5"
FT variation replace (4246,A)
FT /*tag= i
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace (4307,C)
```

```
FT /*tag= j
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace (4403,T)
FT /*tag= k
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace (4628,G)
FT /*tag= l
FT /*standard_name= "Single nucleotide polymorphism"
FT variation replace (4750,A)
FT /*tag= m
FT /*standard_name= "Single nucleotide polymorphism"
XX
PN W0200177125-A2.
XX
PD 18-OCT-2001.
XX
PF 04-APR-2001; 2001WO-US10708.
XX
PR 05-APR-2000; 2000US-194361P.
XX
PA (GENA-) GENAISSANCE PHARM INC.
XX
PI Choi JY, Klieem SE, Koshy B;
XX
DR WPI; 2002-041282/05.
DR P-PSDB; ABH08343.
XX
PT New haplotypes of the human chemokine (C-C motif) receptor 5 gene,
PT useful to diagnose and treat diseases associated with its abnormal
PT expression or function, including human immunodeficiency virus-1
PT infection.
XX
PS Claim 21; Fig 1; 61pp; English.
XX
CC The present sequence represents the human chemokine (C-C motif) receptor
CC 5 (CCR5) gene of the invention which encodes the polypeptide given in
CC AB080343. The specification describes haplotyping the CCR5 gene of an
CC individual by determining if the individual has one of the CCR5
CC haplotypes or haplotype pairs fully defined in the specification. The
CC specification also describes an isolated polynucleotide comprising a
CC nucleotide sequence which is a polymorphic variant of the reference CCR5
CC gene sequence and comprises an isogene defined by a haplotype described
CC in the specification and its encoded polypeptide. The methods of the
CC invention are useful to diagnose and develop treatment for diseases
CC associated with abnormal expression or function of the gene. The CCR5
CC isogenes and the screened compounds are useful for treating human
CC immunodeficiency virus (HIV)-1 infection and the progression to acquired
CC immunodeficiency syndrome (AIDS). The invention has antiviral
CC applications. The specification describes genotyping the CCR5 gene of an
CC individual; predicting a haplotype pair for the CCR5 gene of an
CC individual; identifying an association between a trait and a haplotype or
CC haplotype pair of the CCR5 gene. The specification describes a
CC composition comprising a genotyping oligonucleotide for detecting a CCR5
CC polymorphism; a recombinant non-human organism transformed with CCR5
CC polynucleotide expressing a CCR5 protein encoded by the variant sequence;
CC an isolated antibody specific for the CCR5 polypeptide and a method for
CC screening drugs targeting the CCR5 polypeptide.
XX
SQ Sequence 9141 BP; 2596 A; 1853 C; 2065 G; 2627 T; 0 other;

Query Match 95.4%; Score 1375.2; DB 24; Length 9141;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 1432; Conservative 0; Mismatches 3; Indels 35; Gaps 2;

QY 1 GAATTCCTCCCAACAGAGCCAGCTCTCCATCTAGTGGACAGGGAAGCTAGCAACACCT 60
    |||
Db 3844 GAATTCCTCCCAACAGAGCCAGCTCTCCATCTAGTGGACAGGGAAGCTAGCAACACCT 3903
    |||
QY 61 TCCCTTCACTACAAACTTCATTGCTTGGCCAAAAGAGAGTTAATTCATGTAGCATC 120
    |||
Db 3904 TCCCTTCACTACAAACTTCATTGCTTGGCCAAAAGAGAGTTAATTCATGTAGCATC 3963
    |||
QY 121 TATGTAGGCAATTAAACACCTATTGATGATATAAACACGTTTGCAATCATGGAGGCAACT 180
```

Db 3964 TATGTAGGCAATTAATAAACCTATTGATGATATAAACAGTTTGCAATTCATGGAGGCAACT 4023  
QY 181 AAATACATCTTAGGACTTTAAAGATCACTATTTTATATGACAGAGGTTGGAACAAGA 240  
Db 4024 AAATACATCTTAGGACTTTAAAGATCACTATTTTATATGACAGAGGTTGGAACAAGA 4083  
QY 241 TGGATTATCAAGTGTCAAGTCCAATCTATGACATCAATATATATACATCGGAGCCCTGCC 300  
Db 4084 TGGATTATCAAGTGTCAAGTCCAATCTATGACATCAATATATATACATCGGAGCCCTGCC 4143  
QY 301 AAAAATCAATGTGAAGCAAAATGCGAGCCCGCTCTGCGCTCCGCTCTACTCACTGGTGT 360  
Db 4144 AAAAATCAATGTGAAGCAAAATGCGAGCCCGCTCTGCGCTCCGCTCTACTCACTGGTGT 4203  
QY 361 TCATCTTTGGTTTTTGGGCAACATGCTGTCATCCATCCCTGATTAACACTGCAAAAGGC 420  
Db 4204 TCATCTTTGGTTTTTGGGCAACATGCTGTCATCCCTGATTAACACTGCAAAAGGC 4263  
QY 421 TGAAGACATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTTTCCTTC 480  
Db 4264 TGAAGACATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTTTCCTTC 4323  
QY 481 TTACTGTCCTCTCTGGCTCACTACTGCTGCGCCGCGGAGTGGGACTTTGGAATACAAATGT 540  
Db 4324 TTACTGTCCTCTCTGGCTCACTACTGCTGCGCCGCGGAGTGGGACTTTGGAATACAAATGT 4383  
QY 541 GTCACCTCTTGACGGCTCTATTTATAGGCTTCTCTGGAATCTCTTCATCATCC 600  
Db 4384 GTCACCTCTTGACGGCTCTATTTATAGGCTTCTCTGGAATCTCTTCATCATCC 4443  
QY 601 TCCGACATTCGATGAGTACCTGGCTGCTGCTCCAGCTGTTGTTGTTTAAAGCCAGGA 660  
Db 4444 TCCGACATTCGATGAGTACCTGGCTGCTGCTCCAGCTGTTGTTGTTTAAAGCCAGGA 4503  
QY 661 CGGTCACTTTGGGTTGGTGACAGTGTGATCACTTGGGTTGGTGTTGTTGCTCTC 720  
Db 4504 CGGTCACTTTGGGTTGGTGACAGTGTGATCACTTGGGTTGGTGTTGTTGCTCTC 4563  
QY 721 TCCAGGAATCATCTTTACAGATCTCAAAAGAGTGTCTTATACACTGCGAGTCTC 780  
Db 4564 TCCAGGAATCATCTTTACAGATCTCAAAAGAGTGTCTTATACACTGCGAGTCTC 4623  
QY 781 ATTTTCCAT-----ACATTAAGATAGTCACT 808  
Db 4624 ATTTTCCATGAGTCACTCACTGATCACTTCCAGACATTTAAAGATAGTCACT 4683  
QY 809 TGGGGTGTGCTGCGCTGTTGTCATGTCATCTCTACTCGGGAATCTCTAAAACTC 868  
Db 4684 TGGGGTGTGCTGCGCTGTTGTCATGTCATCTCTACTCGGGAATCTCTAAAACTC 4743  
QY 869 TGCTTGGTGTGGAATCAGAGAGAGGACGACAGGCTGTGAGGCTTATCTTACCATCA 928  
Db 4744 TGCTTGGTGTGGAATCAGAGAGAGGACGACAGGCTGTGAGGCTTATCTTACCATCA 4803  
QY 929 TGATTTGTTATTTCTCTTGGGCTCCCTACACATTTGCTCTCTCTGAAACACCTTC 988  
Db 4804 TGATTTGTTATTTCTCTCTTGGGCTCCCTACACATTTGCTCTCTCTGAAACCTTC 4863  
QY 989 AGGAATCTTTGGCTGGAATATGACAGTGTCTACAGTTGGACCAAGCTATGACGG 1048  
Db 4864 AGGAATCTTTGGCTGGAATATGACAGTGTCTACAGTTGGACCAAGCTATGACGG 4923  
QY 1049 TGACAGAGACTCTTGGGATGACGACTGCTGATCAACCCATCATCTATGCTTTGTCG 1108  
Db 4924 TGACAGAGACTCTTGGGATGACGACTGCTGATCAACCCATCATCTATGCTTTGTCG 4983  
QY 1109 GGGAGAGTTTCAAGACTTACCTCTTCTTCCAAAAGACATTTGCCAAAGCTTCT 1168  
Db 4984 GGGAGAGTTTCAAGACTTACCTCTTCTTCCAAAAGACATTTGCCAAAGCTTCT 5043  
QY 1169 GCAATGCTGTTCTATTTTCCAGCAAGAGGCTCCGAGGAGCAAGCTCAGTTTACACCC 1228

Db 5044 GCAAAATGCTGTTCTATTTTCCAGCAAGAGGCTCCCGAGCGAGCAAGCTCAGTTTACACCC 5103  
QY 1229 GATCCACTGGGGAGCAGGAATATCTGTGGGCTGTGTGACACGGACTCAAGTGGGCTGGTG 1288  
Db 5104 GATCCACTGGGGAGCAGGAATATCTGTGGGCTGTGTGACACGGACTCAAGTGGGCTGGTG 5163  
QY 1289 ACCAGTCCAGAGTGTGACATGCTTGTAGTTTTCATACACAGCCTGGGCTGGGGGTGG-- 1346  
Db 5164 ACCAGTCCAGAGTGTGACATGCTTGTAGTTTTCATACACAGCCTGGGCTGGGGGTGGGG 5223  
QY 1347 -TTGGAGGCTCTTTTAAAGGAAGTACTGTTATAGAGGCTCTAAGATTTCATCCATT 1405  
Db 5224 TGGGAGAGGCTCTTTTAAAGGAAGTACTGTTATAGAGGCTCTAAGATTTCATCCATT 5283  
QY 1406 ATTTGGCATCTGTTTAAAGTAGATTATC 1435  
Db 5284 ATTTGGCATCTGTTTAAAGTAGATTATC 5313  
RESULT 5  
AAF21105  
ID AAF21105 standard; DNA; 143068 BP.  
XX AC AAF21105;  
XX DT 14-MAR-2001 (first entry)  
XX Human low adenosine antisense oligonucleotide related sequence #2672.  
DE Low adenosine antisense oligonucleotide; phosphorothioate; allergy;  
XX human; airway disorder; bronchoconstriction; lung inflammation;  
KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;  
KW immunosuppressive; antisthmatic; analgesic; hypotensive; cytostatic;  
KW respiratory obstruction; pulmonary obstruction; impeded respiration;  
KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;  
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;  
KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;  
KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;  
KW cancer; ss.  
XX Homo sapiens.  
XX WO200062736-A2.  
XX 26-OCT-2000.  
XX 24-MAR-2000; 2000WO-US08020.  
XX 06-APR-1999; 99US-0127958.  
XX (UYEC-) UNIV EAST CAROLINA.  
XX (NYCE/) NYCE J W.  
XX Nyce JW;  
XX WPI; 2000-679539/66.  
XX Low adenosine (A) content antisense oligonucleotides which do not  
PT trigger adenosine receptors during metabolism, useful e.g. for treating  
PT cancers and respiratory obstructions -  
XX  
PS Disclosure; Page 924-957; 1592pp; English.  
XX The present invention describes low adenosine (A) content antisense  
CC oligonucleotides and compositions (I) comprising them. In the antisense  
CC oligonucleotides the A is replaced by a 'Universal' or alternative base.  
CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,  
CC immunosuppressive, antisthmatic, hypotensive and cytostatic activities.  
CC The antisense oligonucleotides and (I) can be used to down-regulate the  
CC expression and or activity of target polypeptides associated with  
CC lung/respiratory disorders and malignancies, such as stimulating and  
CC activating peptide factors and transmitters, transcription factors,  
CC immunoglobulins and antibodies, antibody receptors, cytokines and













CC	antiesene oligonucleotide (ON) with low adenosine (up to 15%), which	CC
CC	targets nucleic acids involved in bronchoconstriction, allergies, and/or	CC
CC	inflammation. The ON can have antiinflammatory, allergies, and/or	CC
CC	antiallergic, cytostatic and analgesic activities. The compositions are	CC
CC	useful for the treatment of diseases associated with inflammation,	CC
CC	impaired airways, including lung disease and diseases whose secondary	CC
CC	effects afflict the lungs of a subject. They can be used for treating	CC
CC	e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,	CC
CC	asthma, impeded respiration, respiratory distress syndrome, pain, cystic	CC
CC	fibrosis, pulmonary hypertension, emphysema, chronic obstructive	CC
CC	pulmonary disease (COPD), and cancers such as leukemias, lymphomas,	CC
CC	carcinomas, and cancers which may metastasize to the lungs, including	CC
CC	breast and prostate cancer. The reduction of the adenosine content of	CC
CC	the ONs reduces side effects. The A-containing ONs break down with the	CC
CC	release of deoxyadenosine which activates adenosine receptors causing	CC
CC	bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the	CC
CC	nucleotide sequences given in the sequence listing from the present	CC
CC	invention, which correspond to SEQ ID NO:1 to 1815, and then the last	CC
CC	185 sequences are also called SEQ ID NO:1 to 185, but the sequences	CC
CC	differ from the previously named sequences. SEQ ID NO:11 to 1680	CC
CC	(AAA32323 to AAA33992) are specifically claimed ONs from the present	CC
CC	invention. N.B. Sequences given in the disclosure of the present	CC
CC	invention do not match up with their corresponding SEQ ID NO: sequences	CC
CC	given in the sequence listing.	CC
XX		CC
SS	Sequence 149412 BP: 43049 A; 31388 C; 33852 G; 41123 T; 0 other;	CC
SEQ		CC
Query Match	95.4%; Score 1375.2; DB 21; Length 149412;	CC
Best Local Similarity	97.4%; Pred. No. 0;	CC
Matches 1432; Conservative	0; Mismatches 3; Indels 35; Gaps 2;	CC
QY	1 GAATTCCTCCCAACAGCGCAAGCTCTCCATCTAGTGGACAGGGAAGCTAGCAGCAAACT 60	CC
Db	67588 GAATTCCTCCCAACAGCGCAAGCTCTCCATCTAGTGGACAGGGAAGCTAGCAGCAAACT 67647	CC
QY	61 TCCTTCACCTACAAACTTCATTCGTGGCCAAAAGAGAGCTTAATTCAAATGTAGATC 120	CC
Db	67648 TCCTTCACCTACAAACTTCATTCGTGGCCAAAAGAGAGCTTAATTCAAATGTAGATC 67700	CC
QY	121 TATGTAGGCAATTAATAAACCCTATTCATCTATATAACAGTTTGCAATTCATGGAGGCAACT 180	CC
Db	67708 TATGTAGGCAATTAATAAACCCTATTCATCTATATAACAGTTTGCAATTCATGGAGGCAACT 67767	CC
QY	181 AATAACATCTAGGACTTTATTAAGATACACTTTTTATTATATGCACAGGGTGGAAACA 240	CC
Db	67768 AATAACATCTAGGACTTTATTAAGATACACTTTTTATTATATGCACAGGGTGGAAACA 67827	CC
QY	241 TGGATTATCAAGTGTCAAGTCCAACTATGACATCAATTTATATACATCGGAGCCCTGCC 300	CC
Db	67828 TGGATTATCAAGTGTCAAGTCCAACTATGACATCAATTTATATACATCGGAGCCCTGCC 67887	CC
QY	301 AAAAATCAATGTGAAGCAAAATCGACAGCCGCCCTCTCGCTCGCTCTACTACTCTGGTGT 360	CC
Db	67888 AAAAATCAATGTGAAGCAAAATCGACAGCCGCCCTCTCGCTCGCTCTACTACTCTGGTGT 67947	CC
QY	361 TCATCTTTGGTTTGTGGGCAACATGCTGFTCATCTCATCTCTGATTAACATGCAAAAGC 420	CC
Db	67948 TCAUCTTTGGTTTGTGGGCAACATGCTGFTCATCTCATCTCTGATTAACATGCAAAAGC 68007	CC
QY	421 TGAAGAGCATGACTGACATCTAGCTGCTCAACCTGGCCATCTCTGACCTGTGTTTTCCCTC 480	CC
Db	68008 TGAAGAGCATGACTGACATCTAGCTGCTCAACCTGGCCATCTCTGACCTGTGTTTTCCCTC 68067	CC
QY	481 TTACTTGTCCTCTCTGGCTGCATATGCTGCCGCCCAACGAGGACTTTGGAATACAATGT 540	CC
Db	68068 TTACTTGTCCTCTCTGGCTGCATATGCTGCCGCCCAACGAGGACTTTGGAATACAATGT 68127	CC
QY	541 GTCAACTCTTCAGAGGCTCTATTTTATAGGCTTCTTCTCTGGAATCTCTTCTTCATATCC 600	CC
Db	68128 GTCAACTCTTCAGAGGCTCTATTTTATAGGCTTCTTCTCTGGAATCTTCTTCTTCATATCC 68187	CC
QY	601 TCCTGACAATCGATAGTACCTGGCTGTCTGCCATCTGCTGTGTTGCTTTTAAAGCCAGGA 660	CC





QY 1169 GCAAAAGCTGTCTATTTTCCAGCAGAGAGGCTCCCGAGCGAGCAAGCTCAGTTTACACC 1228  
DB 68788 GCAAAAGCTGTCTATTTTCCAGCAGAGAGGCTCCCGAGCGAGCAAGCTCAGTTTACACC 68847  
QY 1229 GATCCACTGGGGAGCAGGAATATCTGTGGCTGTGTGACAGGACTCAAGTGGCTGGTG 1288  
DB 68848 GATCCACTGGGGAGCAGGAATATCTGTGGCTGTGTGACAGGACTCAAGTGGCTGGTG 68907  
QY 1289 ACCAGTCAGAGTGTGCACATGCTAGTTTTCATACACAGCCCTGGCTGGGGTGG-- 1346  
DB 68908 ACCAGTCAGAGTGTGCACATGCTAGTTTTCATACACAGCCCTGGCTGGGGTGGGG 68967  
QY 1347 -TTGGAGGCTCTTTTAAAGGAAGTACTGTTATAGAGGCTTAGATTCATCCATT 1405  
DB 68968 TGGGAGAGGCTCTTTTAAAGGAAGTACTGTTATAGAGGCTTAGATTCATCCATT 69027  
QY 1406 ATTTGGCATCTGTTTAAAGTAGATTAGATC 1435  
DB 69028 ATTTGGCATCTGTTTAAAGTAGATTAGATC 69057  
RESULT 11  
AAA34983  
ID AAA34983 standard; DNA; 143068 BP.  
AC AAA34983;  
XX  
DT 28-JUL-2000 (first entry)  
DE Human adenosine receptor related polynucleotide SEQ ID NO:2672.  
KW Human; adenosine receptor; low adenosine antisense oligonucleotide;  
KW phosphothioate; impaired respiration; inflammation; allergy;  
KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;  
KW antiallergic; antiaesthetic; cyostatic; analgesic; impaired airway;  
KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;  
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;  
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;  
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.  
XX Homo sapiens.  
OS  
XX WO200009525-A2.  
PN  
XX  
PD 24-FEB-2000.  
PD  
PF 03-AUG-1999; 99WO-US17712.  
PF  
XX 03-AUG-1998; 98US-0095212.  
PR  
XX (UYEC-) UNIV EAST CAROLINA.  
PA  
XX NYCE JW;  
PI  
XX WPI; 2000-205971/18.  
DR  
XX  
PT New antisense oligonucleotides useful for treating e.g. pulmonary  
PT vasoconstriction, inflammation, allergies, asthma, hypertension  
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or  
PT cancers -  
XX  
PS Disclosure; Page 851-882; 1343pp; English.  
PS  
CC The present invention describes a new composition comprising an  
CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which  
CC targets nucleic acids involved in bronchoconstriction, allergies, and/or  
CC inflammation. The ON can have antiinflammatory, antiallergic,  
CC antiaesthetic, cytostatic and analgesic activities. The compositions are  
CC useful for the treatment of diseases associated with inflammation,  
CC impaired airways, including lung disease and diseases whose secondary  
CC effects afflict the lungs of a subject. They can be used for treating  
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,  
CC asthma, impeded respiration, respiratory distress syndrome, pain, cystic

CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive  
CC pulmonary disease (COPD), and cancers such as leukemias, lymphomas,  
CC carcinomas, and cancers which may metastasize to the lungs, including  
CC breast and prostate cancer. The reduction of the adenosine content of  
CC the ONs reduces side effects. The A-containing ONs break down with the  
CC release of deoxyadenosine which activates adenosine receptors causing  
CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the  
CC nucleotide sequences given in the sequence listing from the present  
CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last  
CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences  
CC differ from the previously named sequences. SEQ ID NO:11 to 1680  
CC (AAA32323 to AAA33992) are specifically claimed ONs from the present  
CC invention. N.B. Sequences given in the disclosure of the present  
CC invention do not match up with their corresponding SEQ ID NO: sequences  
CC given in the sequence listing.  
XX  
SQ Sequence 143068 BP; 41194 A; 30122 C; 32402 G; 39350 T; 0 other;  
Query Match 95.38; Score 1373.6; DB 21; Length 143068;  
Best local Similarity 97.38; Pred. No. 0;  
Matches 1431; Conservative 0; Mismatches 4; Indels 35; Gaps 2;  
QY 1 GAATTCCTCCCAACAGAGCCCAAGCTCTCCATCTAGTGGACAGGAGGCTAGCAGCAAACT 60  
DB 61244 GAATTCCTCCCAACAGAGCCCAAGCTCTCCATCTAGTGGACAGGAGGCTAGCAGCAAACT 61303  
QY 61 TCCTTCTACTACAAACTTCATTCGTTGGCCAAAAGAGAGTTAATTCATGTAGACATC 120  
DB 61304 TCCTTCTACTACAAACTTCATTCGTTGGCCAAAAGAGAGTTAATTCATGTAGACATC 61363  
QY 121 TATGTAGGCAATATAAACCCTTATGATGTATAAAGAGTTTCATTCATGGAGGCACT 180  
DB 61364 TATGTAGGCAATATAAACCCTTATGATGTATAAAGAGTTTCATTCATGGAGGCACT 61423  
QY 181 AAATACATCTAGGAGCTTTATAAAGATCACTTTTATATGACAGAGGTGGAACAAGA 240  
DB 61424 AAATACATCTAGGAGCTTTATAAAGATCACTTTTATATGACAGAGGTGGAACAAGA 61483  
QY 241 TGATTTATCAGTCTCAAGTCCCAATCTATGACATCAATATTATATACATCGAGCCCTGCC 300  
DB 61484 TGGATTATCAGTCTCAAGTCCCAATCTATGACATCAATATTATATACATCGAGCCCTGCC 61543  
QY 301 AAAAAATCAATGTGAAGCAAAATCGAGCCGCCCTCCCTGCTCCGCTCTACTCTGGTGT 360  
DB 61544 AAAAAATCAATGTGAAGCAAAATCGAGCCGCCCTCCCTGCTCCGCTCTACTCTGGTGT 61603  
QY 361 TCATCTTTGGTTTGTGGGCAACATGCTGGTCACTCTCATCTCATGATAACTGCAAAAGGC 420  
DB 61604 TCATCTTTGGTTTGTGGGCAACATGCTGGTCACTCTCATCTCATGATAACTGCAAAAGGC 61663  
QY 421 TGAAGAGCATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTCTCTTC 480  
DB 61664 TGAAGAGCATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTCTCTTC 61723  
QY 481 TTACTGCTCCCTTCTGGGCTCACTATGCTGCGCCGAGCTGGAGCTTTGGAATACAATGT 540  
DB 61724 TTACTGCTCCCTTCTGGGCTCACTATGCTGCGCCGAGCTGGAGCTTTGGAATACAATGT 61783  
QY 541 GTCACATCTTGACAGGCTCTAATTTATAGGCTTTCTCTCGAATCTCTTCATCATCC 600  
DB 61784 GTCACATCTTGACAGGCTCTAATTTATAGGCTTTCTCTCGAATCTCTTCATCATCC 61843  
QY 601 TCCTGACATGATGAGTACCTGGCTGCTGCCATCTGTGTTGCTTTAAAGCCAGGA 660  
DB 61844 TCCTGACATGATGAGTACCTGGCTGCTGCCATCTGTGTTGCTTTAAAGCCAGGA 61903  
QY 661 CGGTACCTTTGGGCTGGTGACAAAGTGTGATCACTTTGGGTGGTGGTGTGTTGGCTTC 720  
DB 61904 CGGTACCTTTGGGCTGGTGACAAAGTGTGATCACTTTGGGTGGTGGTGTGTTGGCTTC 61963  
QY 721 TCCAGGAATCATCTTTACAGATCTCAAAAGAGGTCTTCATTCACCTGACGTCTC 780  
DB 61964 TCCAGGAATCATCTTTACAGATCTCAAAAGAGGTCTTCATTCACCTGACGTCTC 62023





QY 481 TTACTGTCCCTTCCTGGGCTCACTATGCTGCGCCGAGTGGGACTTTGGAAATACAAATGT 540  
Db 481 TTACTGTCCCTTCCTGGGCTCACTATGCTGCGCCGAGTGGGACTTTGGAAATACAAATGT 540  
QY 541 GTCAACTCTTGACAGGCTCTATTTATAGGCTCTCTCTGGAATCTCTTCATCATCC 600  
Db 541 GTCAACTCTTGACAGGCTCTATTTATAGGCTCTCTCTGGAATCTCTTCATCATCC 600  
QY 601 TCCGTGAATCATGATAGTACCTGGCTGCTGCTCCATGCTGTGTTTAAAGCCAGGA 660  
Db 601 TCCGTGAATCATGATAGTACCTGGCTGCTGCTCCATGCTGTGTTTAAAGCCAGGA 660  
QY 661 CGGTCACTTTGGGCTGGTGAAGTGTGATCATCTGGGTGGTGGCTGTGTTGCGTCTC 720  
Db 661 CGGTCACTTTGGGCTGGTGAAGTGTGATCATCTGGGTGGTGGCTGTGTTGCGTCTC 720  
QY 721 TCCAGGATCATCTTTACAGATCTCAAAAGAGGCTCTCATACACCTGAGCTCTC 780  
Db 721 TCCAGGATCATCTTTACAGATCTCAAAAGAGGCTCTCATACACCTGAGCTCTC 780  
QY 781 ATTTTCATACATTAAGATAGTATCTTTGGGCTGCTGCTGCGCTGCTTGTATGCTC 840  
Db 781 ATTTTCATACATTAAGATAGTATCTTTGGGCTGCTGCTGCGCTGCTTGTATGCTC 840  
QY 841 ATCTGTCTCTGGGAACTCTAAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
Db 841 ATCTGTCTCTGGGAACTCTAAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
QY 901 AGGCTGTGAGGCTTATCTTACCATCATGATTTTATTTCTCTCTGCGCTCCCTAC 960  
Db 901 AGGCTGTGAGGCTTATCTTACCATCATGATTTTATTTCTCTCTGCGCTCCCTAC 960  
QY 961 AACATGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
Db 961 AACATGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
QY 1021 TCTAACAGGTTGGACCAAGCTATGCAAGTGTGACAGAGCTCTTGGGATGACGCTGCTG 1080  
Db 1021 TCTAACAGGTTGGACCAAGCTATGCAAGTGTGACAGAGCTCTTGGGATGACGCTGCTG 1080  
QY 1081 ATCAACCCATCATCTATGCTTGTGCGGGAGAGTTCAGAACTACCTCTTATGCTCTC 1140  
Db 1081 ATCAACCCATCATCTATGCTTGTGCGGGAGAGTTCAGAACTACCTCTTATGCTCTC 1140  
QY 1141 TTCCAAAGCATTTGCCAAGCTCTGCAATGCTGCTTATTTTCCACAGAGGCT 1200  
Db 1141 TTCCAAAGCATTTGCCAAGCTCTGCAATGCTGCTTATTTTCCACAGAGGCT 1200  
QY 1201 CCCGAGGAGCAAGCTCAGTTTACCCGATCCACTGGGAGCAGGAAATATCTGTGGG 1260  
Db 1201 CCCGAGGAGCAAGCTCAGTTTACCCGATCCACTGGGAGCAGGAAATATCTGTGGG 1260  
QY 1261 TTGTGACAGGACTCAAGTGGCTGTGACCCAGTGTGATGCTGATGCTGCTGCTGCTGCT 1320  
Db 1261 TTGTGACAGGACTCAAGTGGCTGTGACCCAGTGTGATGCTGATGCTGCTGCTGCTGCT 1320  
QY 1321 TCATACACAGCTGGGCTGGGCT 1344  
Db 1321 TCATACACAGCTGGGCTGGGCT 1344

RESULT 13  
ID AAH26903  
XX AAH26903 standard; cDNA; 1376 BP.  
XX AAH26903;  
AC AAH26903;  
DT 21-DEC-2001 (first entry)  
XX Human HIV-1 co-receptor CCR5 nucleotide sequence.  
XX CCR5; chemokine; co-receptor; human immunodeficiency virus type 1;

KW HIV-1; infection; therapy; vaccine; anti-HIV-1; ss.  
XX Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 240..1298  
XX /\*tag= a  
PN WO200164710-A2.  
XX 07-SEP-2001.  
XX 28-FEB-2001; 2001WO-US06699.  
XX 29-FEB-2000; 2000US-185667P.  
PR 19-MAY-2000; 2000US-205839P.  
PR 07-FEB-2001; 2001US-267231P.  
XX (PROG-) PROGENICS PHARM INC.  
PA (AARO-) ARON DIAMOND AIDS RES CENT.  
XX Dragic T, Olson WC;  
PI WPI; 2001-611273/70.  
DR P-PSDB; AAB82948.  
XX Novel compounds comprising specific amino acids within CCR5 (HIV 1  
PT co-receptor) amino terminal domain including negatively charged and two  
PT sulfated tyrosine residues is useful for treating HIV infection in  
XX humans  
PS Disclosure; Page 30-31; 163pp; English.  
XX The present sequence is that of a polynucleotide encoding the  
CC human HIV-1 co-receptor CCR5 (see AAB82948). Amino acids 2-18  
CC in the N-terminal region of CCR5 comprise an HIV gp120-binding  
CC site that determines the specificity of the interaction between  
CC CCR5 and gp120. Post-translational sulfation of the tyrosine  
CC residues in the CCR5 N-terminus is required for gp120 binding and  
CC may critically modulate the susceptibility of target cells to HIV-1  
CC infection in vivo. The invention provides claimed sulfated  
CC peptides (see AAB82947) that are based on the CCR5 N-terminal  
CC region and which are effective for inhibiting HIV-1 binding to  
CC CCR5. These peptides are used in claimed methods of inhibiting HIV  
CC infection of CD4+ cells, of preventing CD4+ cells from becoming  
CC infected with HIV, of treating a subject whose CD4+ cells are  
CC infected with HIV, and of identifying an agent which inhibits  
CC binding of a CCR5 ligand to a CCR5 receptor. The methods may be  
CC carried out in a subject, especially a human, infected (therapeutic  
CC method), not infected with HIV (prophylactic method), or in a  
CC subject who is not infected with, but has been exposed to, HIV.  
XX Sequence 1376 BP; 347 A; 338 C; 297 G; 394 T; 0 other;  
SQ  
Query Match 90.3%; Score 1302; DB 22; Length 1376;  
Best Local Similarity 97.7%; Pred. No. 0;  
Matches 1344; Conservative 0; Mismatches 0; Indels 32; Gaps 1;  
QY 1 GAATTCCTCCCAACAGAGCCAGCTCTCCATCTAGTGGACAGGAGCTAGCAGCAACCT 60  
Db 1 GAATTCCTCCCAACAGAGCCAGCTCTCCATCTAGTGGACAGGAGCTAGCAGCAACCT 60  
QY 61 TCCCTTCCTCACTACAAAACCTTCATCTGCTGGCCAAAAGAGAGTTAATTCATGTAGCATC 120  
Db 61 TCCCTTCCTCACTACAAAACCTTCATCTGCTGGCCAAAAGAGAGTTAATTCATGTAGCATC 120  
QY 121 TATGTAGGCAATTAATAAACCTTATGATGTATATAAACAGTTTGCATTCATGGAGGCAACT 180  
Db 121 TATGTAGGCAATTAATAAACCTTATGATGTATATAAACAGTTTGCATTCATGGAGGCAACT 180  
QY 181 AATAACATCTTAGGACTTTATAAAGATCACTTTTATTTATGCACAGGTTGGAAACAAGA 240  
Db 181 AATAACATCTTAGGACTTTATAAAGATCACTTTTATTTATGCACAGGTTGGAAACAAGA 240



QY	1	GAATTCCCCACAGAGCCAGCTCTCCATCTAGTGTGGACAGGAGCTAGCAGCAAACT	60
Db	1		
QY	61	TCCCTTCACATACAAAACCTTCATTGCTTGTCGCCAAAGAGAGATTAAATCAATGTAGACATC	120
Db	61		
QY	121	TATGTAGSCAAATTTAAAAACCTTATGTATGTATATAAACAGTTTGCATTCATGGAGGCCAACT	180
Db	121		
QY	181	AAATACATCTAGGACTTATAAAGATCACTTTTATTATTATGACAGAGGTGGAACAGA	240
Db	181		
QY	241	TGGATTATCAAGTGTCAAGTCCAACTATATGACATCAATATTATACATCGGAGCCCTGCC	300
Db	241		
QY	301	AAAAATCAATGTGAAGCAAAATCGAGCCCGCTCTGCGCTCCGCTCTACTCAGTGGTGT	360
Db	301		
QY	361	TCATCTTTGGTTTGTGGGCCAACAATGCTGGTCATCTCTGATATAAATGCAAAAGGC	420
Db	361		
QY	421	TGAAGAGATGACTGACATCTACTGCTCAACCTGGGCCATCTCTGACCTGTTTCCCTTC	480
Db	421		
QY	481	TTACTGTCCCTCTCTGGGCTCACCTATGCTGGCGGCCAGTGGGACTTTGGAATACAAATGT	540
Db	481		
QY	541	GTCAACTCTTGACAGGGCTCTATTATTATAGGCTCTCTCTGTGSAATCTTCTTCATCATCC	600
Db	541		
QY	601	TCCTGACAATCGATAGTACTGGCTGTGTCCTCATGCTGTGTTGCTTTAAAGACGAGGA	660
Db	601		
QY	661	CGGTCAACCTTTGGGGTGGTGAAAGTGTATCACTTTGGGTGGCTGTGTTGGGTCTC	720
Db	661		
QY	721	TCCAGGAATCACTTTACAGATCTCAAAAGAGAGCTCTTCATTACACCTGCAGCTCTC	780
Db	721		
QY	781	ATTTTCCAT-----ACATTTAAGATAGTCACT	808
Db	781		
QY	809	TGGGGTGGTCCCGCTGCTGTATCGTGCATCTGCTACTCGGGATCCTAAAACCTC	868
Db	841	TGGGGTGGTCCCGCTGCTGTATCGTGCATCTGCTACTCGGGATCCTAAAACCTC	900
QY	869	TGCTTCGGTGTGAAATAGAGAGAGGACAGGGCTGTGAGGCTTATCTTCACCATCA	928
Db	901	TGCTTCGGTGTGAAATAGAGAGAGGACAGGGCTGTGAGGCTTATCTTCACCATCA	960
QY	929	TGATTGTATTTCTCTGTTGGGTCCTTACAAATGTCTCTCTCTGAAACCTTCC	988
Db	961	TGATTGTATTTCTCTGTTGGGTCCTTACAAATGTCTCTCTCTGAAACCTTCC	1020
QY	989	AGAAATCTTTGGCTGGAATAATTGCAAGTAGCTCTTAACAGGTTTGGACCAAGCTATGCAGG	1048
Db	1021	AGAAATCTTTGGCTGGAATAATTGCAAGTAGCTCTTAACAGGTTTGGACCAAGCTATGCAGG	1080
QY	1049	TGACAGAGACTTTGGGATGAGGCACCTGCTGCAATCAACCCCATCATCTATGCGCTTTGCG	1108

Search completed: June 8, 2003, 07:47:35  
Job time : 363 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 07:28:02 ; Search time 77 seconds  
(without alignments)  
5743.219 Million cell updates/sec

Title: US-09-938-703-3  
Perfect score: 1442  
Sequence: 1 GAATTCCTCCACAGAGCCCA.....AGTAGTAGTCGGAATTC 1442

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 153338381 residues  
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_NA: \*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/ptodata/1/ina/PCUS\_COMB.seq: \*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1442	100.0	1442	4	US-08-833-752-3
2	1382.8	95.9	1477	4	US-08-833-752-2
3	1344	93.2	1344	3	US-09-087-232A-16
4	1302	90.3	1376	3	US-09-087-232A-12
5	1288.6	89.4	1414	3	US-08-466-343D-1
6	1149	79.7	3383	4	US-08-861-105-13
7	1149	79.7	3383	4	US-08-575-967A-1
8	985	68.3	1059	4	US-08-724-984A-3
9	983.4	68.2	1071	3	US-09-087-232A-14
10	982.8	68.2	5674	4	US-09-293-170-3
11	954.6	66.2	1059	4	US-09-517-605-8
12	792	54.9	792	4	US-08-833-752-1
13	727.4	50.4	2440	4	US-08-724-984A-1
14	683.4	47.4	1979	1	US-08-450-393A-3
15	683.4	47.4	1979	5	US-08-446-669-3
16	683.4	47.4	1979	5	PCT-US95-00476-3
17	612.6	42.5	2232	3	US-08-450-393A-1
18	612.6	42.5	2232	3	US-08-446-669-1
19	612.6	42.5	2232	3	PCT-US95-00476-1
20	385.4	26.7	461	3	US-09-087-232A-11
21	330.2	22.9	2156	1	US-08-012-988A-1
22	319.4	22.1	1915	4	US-08-575-967A-3
23	318.2	22.1	1065	4	US-08-847-296B-2
24	294.8	20.4	1607	3	US-08-875-573-19
25	294.8	20.4	1695	4	US-09-232-878-1
26	242	16.8	2961	4	US-09-177-437-1
27	229.8	15.9	1586	1	US-08-461-244-1

28	194.6	13.5	1161	1	US-08-153-848-31	Sequence 31, Appl
29	194.6	13.5	1161	5	PCT-US93-11153-31	Sequence 31, Appl
30	194.6	13.5	2254	1	US-08-153-848-27	Sequence 27, Appl
31	194.6	13.5	2254	3	US-09-299-843A-27	Sequence 27, Appl
32	194.6	13.5	2254	4	US-09-088-337B-27	Sequence 27, Appl
33	194.6	13.5	2254	5	PCT-US93-11153-27	Sequence 27, Appl
34	194.6	13.5	3119	3	US-09-299-843A-31	Sequence 31, Appl
35	194.6	13.5	3119	4	US-09-088-337B-31	Sequence 31, Appl
36	184.2	12.8	1050	4	US-08-681-192-1	Sequence 1, Appl
37	165	11.4	239	4	US-09-481-288-1	Sequence 1, Appl
38	150.6	10.4	1137	4	US-09-045-583-6	Sequence 6, Appl
39	150.6	10.4	1137	4	US-09-534-185-6	Sequence 6, Appl
40	150.6	10.4	1664	4	US-09-045-583-4	Sequence 4, Appl
41	150.6	10.4	1664	4	US-09-534-185-4	Sequence 4, Appl
42	142.8	9.9	2751	1	US-08-153-848-23	Sequence 23, Appl
43	142.8	9.9	2751	3	US-09-299-843A-23	Sequence 23, Appl
44	142.8	9.9	2751	4	US-09-088-337B-23	Sequence 23, Appl
45	142.8	9.9	2751	5	PCT-US93-11153-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1  
US-08-833-752-3  
; Sequence 3, Application US/08833752  
; Patent No. 6448375  
; GENERAL INFORMATION:  
; APPLICANT: SAMSON, MICHEL  
; APPLICANT: PARMENTIER, MARC  
; APPLICANT: VASSART, GILBERT  
; APPLICANT: LIBERT, FREDERICK  
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR  
; TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobe, Martens, Olson & Bear  
; STREET: 620 Newport Center Drive 16th Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION NUMBER: US/08/833,752  
; FILING DATE: 9-APR-1997  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Altman, Daniel E  
; REGISTRATION NUMBER: 34,115  
; REFERENCE/DOCKET NUMBER:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1442 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 240..884  
; US-08-833-752-3

Query Match 100.0% Score 1442; DB 4; Length 1442;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GAATTCCTCCACAGAGCCAGCTCTCCATCTAGTGGACAGGAGCTAGCAGCAACT 60  
|||||

Db 1 GAATCCCCAACAGAGCAAGCTCTCCATCTAGTGGACAGGGAAGCTAGACCAAACT 60  
QY TCCCTTCACTACAAACTTCAATTCCTGGCCAAAAGAGAGTAAATTCATGTTAGACATC 120  
Db TCCCTTCACTACAAACTTCAATTCCTGGCCAAAAGAGAGTAAATTCATGTTAGACATC 120  
QY TATGTAGGCAATTAAGCACTATGATGATATAAACAAGTTGCAATCATGAGGGCAACT 180  
Db TATGTAGGCAATTAAGCACTATGATGATATAAACAAGTTGCAATCATGAGGGCAACT 180  
QY AAATACATCTAGGACTTTATAAAGATCACTTTTATATGACAGGGTGAACAAGA 240  
Db AAATACATCTAGGACTTTATAAAGATCACTTTTATATGACAGGGTGAACAAGA 240  
QY TGGATATCAAGTGTCAAGTCCAATCATGATGATCAATTAATATATCATCGGAGCCCTGCC 300  
Db TGGATATCAAGTGTCAAGTCCAATCATGATGATCAATTAATATATCATCGGAGCCCTGCC 300  
QY AAAAATCAATGTGAAGCAATCGAGCCGCCCTCTGCTCGCTCTACTACTGCTGCT 360  
Db AAAAATCAATGTGAAGCAATCGAGCCGCCCTCTGCTCGCTCTACTACTGCTGCT 360  
QY TCATCTTTGGTTTTGTGGCAACATGCTGGTCACTCTCTCATCTGATAAATCGCAAAAGCC 420  
Db TCATCTTTGGTTTTGTGGCAACATGCTGGTCACTCTCTCATCTGATAAATCGCAAAAGCC 420  
QY TGAAGAGATGACGACATCTACCTGCTCAACCTGGCCATCTGACCTGTTTTCCTTC 480  
Db TGAAGAGATGACGACATCTACCTGCTCAACCTGGCCATCTGACCTGTTTTCCTTC 480  
QY TTACTGTCCCTTCTGGGCTCACTATGCTGCGCCAGTGGGACTTTGGAAATACAACTG 540  
Db TTACTGTCCCTTCTGGGCTCACTATGCTGCGCCAGTGGGACTTTGGAAATACAACTG 540  
QY GTCAACTCTTGACAGGCTCTATTTATAGGCTTCTCTGGAATCTTCTCATCATCC 600  
Db GTCAACTCTTGACAGGCTCTATTTATAGGCTTCTCTGGAATCTTCTCATCATCC 600  
QY TCCTGACATCATAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
Db TCCTGACATCATAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
QY CGGTCACTTTGGGCTGTGACAAAGTGTGATCACTTGGGTGGTGTGCTGCTGCTGCTC 720  
Db CGGTCACTTTGGGCTGTGACAAAGTGTGATCACTTGGGTGGTGTGCTGCTGCTGCTC 720  
QY TCCAGAGATCATCTTTACAGATCTCAAAAGAGTCTTATATACACCTGAGCTGCTC 780  
Db TCCAGAGATCATCTTTACAGATCTCAAAAGAGTCTTATATACACCTGAGCTGCTC 780  
QY ATTTTGCATACATTAAGATAGTCACTTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
Db ATTTTGCATACATTAAGATAGTCACTTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
QY ATCTGTACTCGGGAATCCFAAAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
Db ATCTGTACTCGGGAATCCFAAAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
QY AGGCTGTGAGGCTTATCTTACCATCATGATTTGTTATTTCTTCTTCTGCTGCTGCTGCT 960  
Db AGGCTGTGAGGCTTATCTTACCATCATGATTTGTTATTTCTTCTTCTGCTGCTGCTGCT 960  
QY AACATGCTCTCTCTGACACCTTCCAGGAATCTTTGGGCTGAATTAATTCAGTAC 1020  
Db AACATGCTCTCTCTGACACCTTCCAGGAATCTTTGGGCTGAATTAATTCAGTAC 1020  
QY TCTAACAGTGTGGAACCAAGCTATGAGGTGACAGACTCTTGGGATGACGACTGCTG 1080  
Db TCTAACAGTGTGGAACCAAGCTATGAGGTGACAGACTCTTGGGATGACGACTGCTG 1080  
QY ATCAACCCCATCATCTATGCTTGTGCGGGAGAGTTTCAGAACTACTCTTTAGTCTTC 1140  
Db ATCAACCCCATCATCTATGCTTGTGCGGGAGAGTTTCAGAACTACTCTTTAGTCTTC 1140

QY 1141 TTCCAAAGCACATTCGCAACGCTTCTCAATCTGTTCTATTTTCCAGCAAGGCT 1200  
Db 1141 TTCCAAAGCACATTCGCAACGCTTCTCAATCTGTTCTATTTTCCAGCAAGGCT 1200  
QY 1201 CCCGAGGAGCAAGCTCAGTTTACCCGATCCATCGGGAGCAGGAATATCTGTGGC 1260  
Db 1201 CCCGAGGAGCAAGCTCAGTTTACCCGATCCATCGGGAGCAGGAATATCTGTGGC 1260  
QY 1261 TTGTGACAGGACTCAAGTGGGTGGTGGCCAGTCCAGTCTGACATGCTGCTGCT 1320  
Db 1261 TTGTGACAGGACTCAAGTGGGTGGTGGCCAGTCCAGTCTGACATGCTGCTGCT 1320  
QY 1321 TCATACACAGCTGCTGGGTGGTGGGTGGTGGGTGGTGGGTGGTGGGTGGTGGT 1380  
Db 1321 TCATACACAGCTGCTGGGTGGTGGGTGGTGGGTGGTGGGTGGTGGGTGGTGGT 1380  
QY 1381 TAGAGGCTCTAAGATTCATTCATTTATTTGGCATCTGTTTAAAGTAGATAGATCCGAAT 1440  
Db 1381 TAGAGGCTCTAAGATTCATTCATTTATTTGGCATCTGTTTAAAGTAGATAGATCCGAAT 1440  
QY 1441 TC 1442  
Db 1441 TC 1442

## RESULT 2

US-08-833-752-2  
; Sequence 2, Application US/08833752  
; Patent No. 6448375  
; GENERAL INFORMATION:  
; APPLICANT: SAMSON, MICHEL  
; APPLICANT: PARMENTIER, MARC  
; APPLICANT: VASSART, GILBERT  
; APPLICANT: LIBERT, FREDERICK  
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR  
; TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 620 Newport Center Drive 16th Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/833,752  
; FILING DATE: 9-APR-1997  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Altman, Daniel E  
; REGISTRATION NUMBER: 34,115  
; REFERENCE/DOCKET NUMBER:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1477 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 240..1295  
US-08-833-752-2

Query Match 95.9%; Score 1382.8; DB 4; Length 1477;  
Best Local Similarity 97.4%; Pred. No. 0;  
Matches 1439; Conservative 0; Mismatches 3; Indels 35; Gaps 2;





LOCATION: 240..887

US-09-087-232A-16

Query Match 93.2%; Score 1344; DB 3; Length 1344;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1344; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCCTCCCAAGAGGAGGCTCCATCTAGTGGACAGGAGGAGCTAGCAGCAACCT 60  
DB 1 GAATTCCTCCCAAGAGGAGGCTCCATCTAGTGGACAGGAGGAGCTAGCAGCAACCT 60  
QY 61 TCCCTTCACTACAAAATTCATCTGCTGGCCAAAAGAGAGTTAAATCAATAGACATC 120  
DB 61 TCCCTTCACTACAAAATTCATCTGCTGGCCAAAAGAGAGTTAAATCAATAGACATC 120  
QY 121 TATGTAGGCAATTAAGCACTTATGATGATATAAAGAGTTGCAATTCATGGAGGCACT 180  
DB 121 TATGTAGGCAATTAAGCACTTATGATGATATAAAGAGTTGCAATTCATGGAGGCACT 180  
QY 181 AAATACATCTTAGGACCTTTATAAAGATCACTTTTATATGACAGAGGTGGAACAGA 240  
DB 181 AAATACATCTTAGGACCTTTATAAAGATCACTTTTATATGACAGAGGTGGAACAGA 240  
QY 241 TGGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATTAATATACATCGGAGCCCTGCC 300  
DB 241 TGGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATTAATATACATCGGAGCCCTGCC 300  
QY 301 AAAAATCAATGTGAAGCAAAATCGAGCCCGCTCTGCTCCGCTCTACTCAGTGGTGT 360  
DB 301 AAAAATCAATGTGAAGCAAAATCGAGCCCGCTCTGCTCCGCTCTACTCAGTGGTGT 360  
QY 361 TCATCTTTGGTTTTGTGGCAACATGCTGCTCATCTCATCTGATAAACTGCAAAAGGC 420  
DB 361 TCATCTTTGGTTTTGTGGCAACATGCTGCTCATCTCATCTGATAAACTGCAAAAGGC 420  
QY 421 TGAAGACATGACTGACATFACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCCCTTC 480  
DB 421 TGAAGACATGACTGACATFACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCCCTTC 480  
QY 481 TTACGTGCTCCCTCTGCGGCTCACTGCTGCGCCCACTGGGACTTGGAAATACAATGT 540  
DB 481 TTACGTGCTCCCTCTGCGGCTCACTGCTGCGCCCACTGGGACTTGGAAATACAATGT 540  
QY 541 GTCAACTCTTGACAGGCTCTATTTATAGGCTTTCTCTGGAATCTTCTTCATCATCC 600  
DB 541 GTCAACTCTTGACAGGCTCTATTTATAGGCTTTCTCTGGAATCTTCTTCATCATCC 600  
QY 601 TCTTGACATGATAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
DB 601 TCTTGACATGATAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
QY 661 CGGTCACTTTGGGTGGTGAAGTGTGATCACTTGGGTGGTGGTGGTGGTGGTGGTGGTGC 720  
DB 661 CGGTCACTTTGGGTGGTGAAGTGTGATCACTTGGGTGGTGGTGGTGGTGGTGGTGGTGC 720  
QY 721 TCCAGGATCATCTTTACAGATCTCAAAAAGAGGCTTCAATACACTCGACCTC 780  
DB 721 TCCAGGATCATCTTTACAGATCTCAAAAAGAGGCTTCAATACACTCGACCTC 780  
QY 781 ATTTTCCATCATTAAGATGATCATCTTGGGCTGGTCTGCGGCTGCTGCTGCTGCTGCTGCT 840  
DB 781 ATTTTCCATCATTAAGATGATCATCTTGGGCTGGTCTGCGGCTGCTGCTGCTGCTGCTGCT 840  
QY 841 ATCTGCTACTCGGGAATCTTAAAACTCTGCTTGGTGTGCAAAATGAGAGAGGAGGCAC 900  
DB 841 ATCTGCTACTCGGGAATCTTAAAACTCTGCTTGGTGTGCAAAATGAGAGAGGAGGCAC 900  
QY 901 AGGGCTGTAGGCTTATCTCACCACATCATGATGTTATTTCTCTCTGCGCTCCCTTAC 960  
DB 901 AGGGCTGTAGGCTTATCTCACCACATCATGATGTTATTTCTCTCTGCGCTCCCTTAC 960  
QY 961 ACATGTCCTCTCTGACACCTTCCAGAACTTCTGGCTGAAATATTTGCAATGAGTAC 1020  
DB 961 ACATGTCCTCTCTGACACCTTCCAGAACTTCTGGCTGAAATATTTGCAATGAGTAC 1020

DB 961 AACATGTCCTCTCTGACACCTTCCAGGAATCTTTGGCTGAAATATTTGCAATGAGTAC 1020  
QY 1021 TCTAACAGGTTGGACCAAGCTATGAGGTGACAGACCTCTTTGGGATGAGCACTGCTGC 1080  
DB 1021 TCTAACAGGTTGGACCAAGCTATGAGGTGACAGACCTCTTTGGGATGAGCACTGCTGC 1080  
QY 1081 ATCAACCCCATCATCTATGCTTTGTCGGGGAAGATTGAGAACTACCTCTTATGCTTC 1140  
DB 1081 ATCAACCCCATCATCTATGCTTTGTCGGGGAAGATTGAGAACTACCTCTTATGCTTC 1140  
QY 1141 TTCCAAAGACATCTGCAAAAGCTTCTGCAAAATGCTGTTATTTTCCAGCAAGAGCT 1200  
DB 1141 TTCCAAAGACATCTGCAAAAGCTTCTGCAAAATGCTGTTATTTTCCAGCAAGAGCT 1200  
QY 1201 CCGAGGAGCAAGCTCAGTTTACACCCGATCCACCTGGGAGCAGGAAATATCTGTGGC 1260  
DB 1201 CCGAGGAGCAAGCTCAGTTTACACCCGATCCACCTGGGAGCAGGAAATATCTGTGGC 1260  
QY 1261 TTGTGACAGGACTCAAGTGGGCTGGTCCACCTGAGAGTTGTGCACATGGCTAGTTT 1320  
DB 1261 TTGTGACAGGACTCAAGTGGGCTGGTCCACCTGAGAGTTGTGCACATGGCTAGTTT 1320  
QY 1321 TCATACACAGCTGGGCTGGGCT 1344  
DB 1321 TCATACACAGCTGGGCTGGGCT 1344

RESULT 4

US-09-087-232A-12  
; Sequence 12, Application US/09087232A  
; Patent No. 6153431  
; GENERAL INFORMATION:  
; APPLICANT: Quillent et al.  
; TITLE OF INVENTION: HUMAN IMMUNODEFICIENCY VIRUS CO-RECEPTOR  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Baker & Botts, L.L.P. attn. Lisa Kole  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/087,232A  
FILING DATE: 28 MAY 1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/048,057  
FILING DATE: 30 MAY 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: KOLE, LISA B.  
REGISTRATION NUMBER: 35,225  
REFERENCE/DOCKET NUMBER: AP 31115  
TELEPHONE: (212) 408-2628  
TELEFAX: (212) 765-2519  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1376 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 240..1298  
US-09-087-232A-12

Query Match 90.38; Score 1302; DB 3; Length 1376;  
Best Local Similarity 97.74; Pred. No. 0;  
Matches 1344; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 1 GAATTCCTCCCAACAGAGCCAAAGCTCTCCATCTAGTGGACAGGAAGCTAGCAGCAAACT 60  
DB 1 GAATTCCTCCCAACAGAGCCAAAGCTCTCCATCTAGTGGACAGGAAGCTAGCAGCAAACT 60

QY 61 TCCCTTCTACTACAAAACCTTCAATGCTTGGCCAAAAAGAGAGTTAATCAATGAGACATC 120  
DB 61 TCCCTTCTACTACAAAACCTTCAATGCTTGGCCAAAAAGAGAGTTAATCAATGAGACATC 120

QY 121 TAGTAGGCAATTAATAACCTATTGATGTATATAACAGTTTGCATTATGGAGGCACT 180  
DB 121 TAGTAGGCAATTAATAACCTATTGATGTATATAACAGTTTGCATTATGGAGGCACT 180

QY 181 AATACATCTAGGACCTTTATAAAGATCACTTTTATTTATGACAGAGGTGGAACAAGA 240  
DB 181 AATACATCTAGGACCTTTATAAAGATCACTTTTATTTATGACAGAGGTGGAACAAGA 240

QY 241 TGGATTATCAAGTCTCAAGTCCAAATCTATGACATCAATTTATTTATACATCGGAGCCCTGCC 300  
DB 241 TGGATTATCAAGTCTCAAGTCCAAATCTATGACATCAATTTATTTATACATCGGAGCCCTGCC 300

QY 301 AAAAAATCAATGTGAAGCAAAATCGCAGCCGCTCTGCTCCGCTCTACTCACTGGTGT 360  
DB 301 AAAAAATCAATGTGAAGCAAAATCGCAGCCGCTCTGCTCCGCTCTACTCACTGGTGT 360

QY 361 TCATCTTTGGTTTTTGGGCAACATGCTGTCTCATCTCTCATCTGATAAAGTCAAAAGGC 420  
DB 361 TCATCTTTGGTTTTTGGGCAACATGCTGTCTCATCTCTCATCTGATAAAGTCAAAAGGC 420

QY 421 TGAAGAGCATGACATGACATCTACCTGTCTCAACCTGCCATCTCTGACCTCTTTTCCTTC 480  
DB 421 TGAAGAGCATGACATGACATCTACCTGTCTCAACCTGCCATCTCTGACCTCTTTTCCTTC 480

QY 481 TTACTGTCCCTCTCTGGGCTCACTATGCTGCCGCGGAGTGGGACTTTGGAAATACAAATGT 540  
DB 481 TTACTGTCCCTCTCTGGGCTCACTATGCTGCCGCGGAGTGGGACTTTGGAAATACAAATGT 540

QY 541 GTCAACTCTTGACAGGSCCTATTTTATPAGGCTTCTCTCTGGAATCTTCTCATCATCC 600  
DB 541 GTCAACTCTTGACAGGSCCTATTTTATPAGGCTTCTCTCTGGAATCTTCTCATCATCC 600

QY 601 TCCGCAATCGATAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
DB 601 TCCGCAATCGATAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660

QY 661 CGGTCACTTTGGGGTGGTGACAAGTGTGATCACTTGGGTGGTGGTGGTGGTGGTGGTGGTGGT 720  
DB 661 CGGTCACTTTGGGGTGGTGACAAGTGTGATCACTTGGGTGGTGGTGGTGGTGGTGGTGGTGGT 720

QY 721 TCCAGGAATCATCTTTACAGATCTCAAAAGAGGTCTTCAATACACCTGACGCTCTC 780  
DB 721 TCCAGGAATCATCTTTACAGATCTCAAAAGAGGTCTTCAATACACCTGACGCTCTC 780

QY 781 ATTTTCCAT-----ACATTAAGATAGTACT 808  
DB 781 ATTTTCCATAGTACTATCAATTTCTGGAAGATTTCCAGACATTAAGATAGTACT 840

QY 809 TGGGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 868  
DB 809 TGGGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 868

QY 869 TGGTTCGGTTCGAAATGAGAGAGAGGACAGGCTGTGAGGCTTATCTTCAACATCA 928  
DB 901 TGGTTCGGTTCGAAATGAGAGAGAGGACAGGCTGTGAGGCTTATCTTCAACATCA 960

QY 929 TGATTTCTTTATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 988  
DB 961 TGATTTCTTTATTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020

QY 989 AGGAATCTTTGGCTGAATAATTCAGTAGTCTTAACAGGTTGGACCAAGCTATGCAGG 1048  
DB 1021 AGGAATCTTTGGCTGAATAATTCAGTAGTCTTAACAGGTTGGACCAAGCTATGCAGG 1080

QY 1049 TGACAGAGACTCTTGGGATGACGACTGCTGATCAACCCCATCATCTATGCTTTGTGG 1108  
DB 1081 TGACAGAGACTCTTGGGATGACGACTGCTGATCAACCCCATCATCTATGCTTTGTGG 1140

QY 1109 GGGAGAAGTTCAGAACTTCTTATTTCCACAGAGGCTCCCGAGCGAGCAAGCTCAGTTTACACC 1228  
DB 1141 GGGAGAAGTTCAGAACTTCTTATTTCCACAGAGGCTCCCGAGCGAGCAAGCTCAGTTTACACC 1260

QY 1169 GCAAAAGCTGTTCTTATTTTCCACAGAGGCTCCCGAGCGAGCAAGCTCAGTTTACACC 1228  
DB 1201 GCAAAAGCTGTTCTTATTTTCCACAGAGGCTCCCGAGCGAGCAAGCTCAGTTTACACC 1260

QY 1229 GATCCACTGGGAGCAGCAAAATATCTGTGGGCTTGTGACACGAGCTCAAGTGGGCTGGT 1288  
DB 1261 GATCCACTGGGAGCAGCAAAATATCTGTGGGCTTGTGACACGAGCTCAAGTGGGCTGGT 1320

QY 1289 ACCCAGTCAAGTTGTGACATGCTTAGTTTATACACAGGCTGGGCTGGGGT 1344  
DB 1321 ACCCAGTCAAGTTGTGACATGCTTAGTTTATACACAGGCTGGGCTGGGGT 1376

RESULT 5  
US-08-466-343D-1  
; Sequence 1, Application US/08466343D  
; Patent No. 6025154  
; GENERAL INFORMATION:  
; APPLICANT: LI, YI  
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING HUMAN G-PROTEIN  
; TITLE OF INVENTION: CHEMOKINE RECEPTOR HDNR10 (AS AMENDED)  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVE., NW, SUITE 600  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/466,343D  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: STEFFE, ERIC K.  
; REGISTRATION NUMBER: 36,688  
; REFERENCE/DOCKET NUMBER: 1488.1150000/EKS/KLM  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1414 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 259..1314  
US-08-466-343D-1

Query Match 89.44; Score 1288.6; DB 3; Length 1414;  
Best Local Similarity 96.3%; Pred. No. 0;  
Matches 1342; Conservative 0; Mismatches 19; Indels 32; Gaps 1;

QY 2 AATTCCTCCCAACAGAGCCAGCTCTCCATCTAGTGGACAGGAGCTAGCAGCAAAACCTT 61  
Db 21 ATTCCCTCCCAACAGAGCCAGCTCTCCATCTAGTGGACAGGAGCTAGCAGCAAAACCTT 80  
QY 62 CCCTTCACCTACCAAACTTCATCTGTCGCAAAAGAGAGTTAATTCATGTAGACATCT 121  
Db 81 CCCTTCACCTACCAAACTTCATCTGTCGCAAAAGAGAGTTAATTCATGTAGACATCT 140  
QY 122 ATGTAGGCAATTAATAACCTTATGATGATATAAACAAGTTTGCATTCATGAGGAGCAATP 181  
Db 141 ATGTAGGCAATTAATAACCTTATGATGATATAAACAAGTTTGCATTCATGAGGAGCAATP 200  
QY 182 AATACATCTAGGACTTATAAAGATCACATTTTATATGCACAGGGTGGAAACAGAT 241  
Db 201 AATACATCTAGGACTTATAAAGATCACATTTTATATGCACAGGGTGGAAACAGAT 260  
QY 242 GGATATCAAGTGTCAAGTCCCAATCTATGACATCAATTAATATACATCGAGCCCTGCCA 301  
Db 261 GGATATCAAGTGTCAAGTCCCAATCTATGACATCAATTAATATACATCGAGCCCTGCCA 320  
QY 302 AAAATCAATGTGAAGCAATTCGAGCCGCTCTCCCTCCGCTCTACTACTACTGTGTGT 361  
Db 321 AAAATCAATGTGAAGCAATTCGAGCCGCTCTCCCTCCGCTCTACTACTACTGTGTGT 380  
QY 362 CATCTTGTGTTGGGCAACATGTGTGTCATCTCTCATCTGATATAAACTGCAAAAGCT 421  
Db 381 CATCTTGTGTTGGGCAACATGTGTGTCATCTCTCATCTGATATAAACTGCAAAAGCT 440  
QY 422 GAAGAGCATGACTGACATCTACCTGCTCAACTGGCCATCTGACCTGTGTTTCCCTCT 481  
Db 441 GAAGAGCATGACTGACATCTACCTGCTCAACTGGCCATCTGACCTGTGTTTCCCTCT 500  
QY 482 TACTGTCCCTCTGCGCTCACTATGCTGCGCCAGTGGAGCTTGGAAATCAATGTG 541  
Db 501 TACTGTCCCTCTGCGCTCACTATGCTGCGCCAGTGGAGCTTGGAAATCAATGTG 560  
QY 542 TCAACTCTTGACAGGGCTCTATTTATAGCTTCTCTGGAATCTCTTCATCATCT 601  
Db 561 TCAACTCTTGACAGGGCTCTATTTATAGCTTCTCTGGAATCTCTTCATCATCT 620  
QY 602 CCTGACATCAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661  
Db 621 CCTGACATCAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 680  
QY 662 GGTCACTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 721  
Db 681 GGTCACTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 740  
QY 722 CCCAGGATCACTTTTACAGATCTCAAAAGAGGCTTCAATACACTGCACTGCTCA 781  
Db 741 CCCAGGATCACTTTTACAGATCTCAAAAGAGGCTTCAATACACTGCACTGCTCA 800  
QY 782 TTTTCCAT-----ACATTAAGATAGTCACTT 809  
Db 801 TTTTCCATAGCTAGTCACTTCAATTTGGAAGATTTCCAGACATTAAGATAGTCACTT 860  
QY 810 GGGGCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 869  
Db 861 GGGGCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920  
QY 870 GCTTCGGTGTGGAATAGAGAGAGAGGACAGGGCTGTGAGGCTTATCTTACCATCAT 929  
Db 921 GCTTCGGTGTGGAATAGAGAGAGAGGACAGGGCTGTGAGGCTTATCTTACCATCAT 980  
QY 930 GATTTGTTATTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 989  
Db 981 GATTTGTTATTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1040  
QY 990 GGAATTTCTTGGCTGAATAATTCAGTAGCTCTTAAAGAGTGGACCAAGCTATGAGGT 1049  
Db 1041 GGAATTTCTTGGCTGAATAATTCAGTAGCTCTTAAAGAGTGGACCAAGCTATGAGGT 1100  
QY 1050 GACAGAGACTTTGGGATGAGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1109

Db 1101 GACAGAGACTTTGGGATGAGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1160  
QY 1110 GGAGAGCTTCAGAACTTACCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1169  
Db 1161 GGAGAGCTTCAGAACTTACCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1220  
QY 1170 CAATGCTGTCTTATTTCCAGCAAGAGGCTCCGAGGAGCAAGCTCAGTTTACACCG 1229  
Db 1221 CAATGCTGTCTTATTTCCAGCAAGAGGCTCCGAGGAGCAAGCTCAGTTTACACCG 1280  
QY 1230 ATCCACTGGGAGGAGAAATATCTGTTGGCTGTGACAGGACTCAAGTGGCTGGTGA 1289  
Db 1281 ATCCACTGGGAGGAGAAATATCTGTTGGCTGTGACAGGACTCAAGTGGCTGGTGA 1340  
QY 1290 CCCAGTCAGATTTGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1349  
Db 1341 CCCAGTCAGATTTGTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1400  
QY 1350 GGAGGCTCTTTT 1362  
Db 1401 GGAAGAGGCTTT 1413  
RESULT 6  
US-08-861-105-13  
; Sequence 13, Application US/08861105  
; Patent No. 6258527  
; GENERAL INFORMATION:  
; APPLICANT: LITTMAN, DAN R.  
; APPLICANT: DENG, HONGKUI  
; APPLICANT: ELMER, WILFRIED  
; APPLICANT: LANDAU, NATHANIEL R.  
; APPLICANT: LIU, RONG  
; TITLE OF INVENTION: G-COUPLED RECEPTORS ASSOCIATED WITH  
; TITLE OF INVENTION: MACROPHAGE-TROPIC HIV, AND DIAGNOSTIC AND THERAPEUTIC  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David A. Jackson, Esq.  
; STREET: 411 Hackensack Ave, Continental Plaza, 4th  
; STREET: Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/861,105  
; FILING DATE:  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/666,020  
; FILING DATE: 19-JUN-1996  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/227,319  
; FILING DATE: 13-APR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 1049-1-004 N1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 201-487-5800  
; TELEFAX: 201-343-1684  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3383 base pairs  
; TYPE: nucleic acid



QY 224 CACAGGTTGGAACAGATGGATATCAAGTGTCAAGTCCCAATCATATGACATCAATTA 283  
DB 39 CCGCGGTGGAACAGATGGATATCAAGTGTCAAGTCCCAATCATATGACATCAATTA 98  
QY 284 TACATCGAGCCCTGCAAAATCAATGTGAAGCAATCGACGCCGCTCTGCTCC 343  
DB 99 TACATCGAGCCCTGCAAAATCAATGTGAAGCAATCGACGCCGCTCTGCTCC 158  
QY 344 GCTCTACTACTGCTGCTCACTTTGCTTTGCTGGAACATGCTGCTCATCTCATCT 403  
DB 159 GCTCTACTACTGCTGCTCACTTTGCTTTGCTGGAACATGCTGCTCATCTCATCT 218  
QY 404 GATTAATGCAAAAGCTGGAAGCATGACTGACATCTACCTGCTCAACTGGCCATCTC 463  
DB 219 GATAAATGCAAAAGCTGGAAGCATGACTGACATCTACCTGCTCAACTGGCCATCTC 278  
QY 464 TGACCTGTTTTCTCTTCTTACTGCTCCCTCTCTGGCTCAGTATGCTGCGGCCAGTGGGA 523  
DB 279 TGACCTGTTTTCTCTTACTGCTCCCTCTCTGGCTCAGTATGCTGCGGCCAGTGGGA 338  
QY 524 CTTTGGAAATCAAGTGTCAACTCTTGAAGGCTCTATTATAGGCTTCTTCTG 583  
DB 339 CTTTGGAAATCAAGTGTCAACTCTTGAAGGCTCTATTATAGGCTTCTTCTG 398  
QY 584 AATCTTCTCATCTCCTCCTGACAATCGATAGTACCTGCTGCTGCTCATCTGCTGTT 643  
DB 399 AATCTTCTCATCTCCTCCTGACAATCGATAGTACCTGCTGCTGCTCATCTGCTGTT 458  
QY 644 TGCTTTAAAGCCAGGAGGTCACCTTTGGGTTGAGTGTGATGATGCTGCTGCTGTT 703  
DB 459 TGCTTTAAAGCCAGGAGGTCACCTTTGGGTTGAGTGTGATGATGCTGCTGCTGTT 518  
QY 704 GCTGCTGTTGCTCTCCAGCAATCATCTTTACGATCTCAAAAGAGGCTCTCA 763  
DB 519 GCTGCTGTTGCTCTCCAGCAATCATCTTTACGATCTCAAAAGAGGCTCTCA 578  
QY 764 TTACACCTGAGCTCTCAATTTTCCAT-----AC 791  
DB 579 TTACACCTGAGCTCTCAATTTTCCATGATGATGATGATGATGATGATGATGATGAT 638  
QY 792 ATTAAGATAGTCACTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 851  
DB 639 ATTAAGATAGTCACTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 698  
QY 852 GGAATCTTAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 911  
DB 699 GGAATCTTAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 758  
QY 912 GCTTATCTTCAACATCATGATGTTTATTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 971  
DB 759 GCTTATCTTCAACATCATGATGTTTATTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 818  
QY 972 TCTCTTGAACATCTTCCAGGATCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1031  
DB 819 TCTCTTGAACATCTTCCAGGATCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 878  
QY 1032 GGACCAAGCTATGAGGTGACAGACTTTGGGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1091  
DB 879 GGACCAAGCTATGAGGTGACAGACTTTGGGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 938  
QY 1092 CACTATGCTTTGCTGGGAGAGTTTCAAGAACTACCTCTTATGCTTCTTCTTCTTCTTCTTCTT 1151  
DB 939 CACTATGCTTTGCTGGGAGAGTTTCAAGAACTACCTCTTATGCTTCTTCTTCTTCTTCTTCTT 998  
QY 1152 CATTGCCAAGCTTCTGCAATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1211  
DB 999 CATTGCCAAGCTTCTGCAATGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1058  
QY 1212 AAGCTCAGTTTACACCGATCCACTGGGAGCAGGAATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1271  
DB 1059 AAGCTCAGTTTACACCGATCCACTGGGAGCAGGAATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1118

QY 1272 ACTCAAGTGGCTGGTGAOCAGTGTGACATGCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1331  
DB 1119 ACTCAAGTGGCTGGTGAOCAGTGTGACATGCTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1178  
QY 1332 CTGGGCTGGGCTGG--TTGGAGGCTCTTTTAAAGGAGTACTGTTATAGAGGT 1388  
DB 1179 CTGGGCTGGGCTGGGAGAGGCTCTTTTAAAGGAGTACTGTTATAGAGGT 1238  
QY 1389 CTAAGATTCATCCATTTATTTGGCATCTGTTTAAAGTATAGATC 1435  
DB 1239 CTAAGATTCATCCATTTATTTGGCATCTGTTTAAAGTATAGATC 1285

RESULT 8  
US-08-724-984A-3  
; Sequence 3, Application US/08724984A  
; Patent No. 6388055  
; GENERAL INFORMATION:  
; APPLICANT: Derk Bergsma, Mary Brawner, and Usman Shabon  
; TITLE OF INVENTION: No. 6388055el Mouse Genomic Clone of the CC-  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Smithkline Beecham Corporation  
; STREET: 709 Swedeland Road, P.O. Box 1539  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19406-0939  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE  
; COMPUTER: IBM 486  
; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS  
; SOFTWARE: MICROSOFT WORD  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/724,984A  
; FILING DATE: October 3, 1996  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: William T. Han  
; REGISTRATION NUMBER: 34,344  
; REFERENCE/DOCKET NUMBER: ATG50023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610 270 5024  
; TELEFAX: 610 270 5090  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1059  
; TYPE: Nucleic Acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Linear  
; ANTI-SENSE: No  
US-08-724-984A-3

Query Match 68.38; Score 985; DB 4; Length 1059;  
Best Local Similarity 97.08; Pred. No. 2e-274;  
Matches 1027; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 240 ATGGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATTTATATACATCGGAGCCCTGC 299  
DB 1 ATGGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATTTATATACATCGGAGCCCTGC 60  
QY 300 CAAAAATCAATGGAAGCAATCGACGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 359  
DB 61 CAAAAATCAATGGAAGCAATCGACGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
QY 360 TTCATCTTGGTTTGGGCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 419  
DB 121 TTCATCTTGGTTTGGGCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180





; Sequence 8, Application US/09517605

; Patent No. 6391567

; GENERAL INFORMATION:

; APPLICANT: Littman, Dan R.

; APPLICANT: Kwon, Douglas S.

; APPLICANT: Van Kooyk, Ivette

; APPLICANT: Gelitenbeck, Theo

; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO

; FILE REFERENCE: 1049-1-017

; CURRENT APPLICATION NUMBER: US/09/517,605

; CURRENT FILING DATE: 2000-03-02

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 8

; LENGTH: 1059

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-09-517-605-8

Query Match 66.2%; Score 954.6; DB 4; Length 1059;

Best Local Similarity 95.2%; Pred. No. 1.2e-265;

Matches 1008; Conservative 0; Mismatches 19; Indels 32; Gaps 1;

QY 240 ATGATTATCAAGTGTCAAGTCCAACTATGACATCAATATTATATACATCGAGCCCTGC 299

DB 1 ATGATTATCAAGTGTCAAGTCCAACTATGACATCAATATTATATACATCGAGCCCTGC 60

QY 300 CAAAAATCAATGTGAAGCAATCGAGCCCGCTCCGCTCGGCTCTACTCTACTGGTG 359

DB 61 CAAAAATCAATGTGAAGCAATCGAGCCCGCTCCGCTCGGCTCTACTCTACTGGTG 120

QY 360 TTCACTCTTTGGTTTGGCAACATGCTGCTCATCTCTCTCACTGATTAACGCAAAAGG 419

DB 121 TTCACTCTTTGGTTTGGCAACATGCTGCTCATCTCTCTCACTGATTAACGCAAAAGG 180

QY 420 CTGAAGAGCATGATGACATCTACTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT 479

DB 181 CTGAAGAGCATGATGACATCTACTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTT 240

QY 480 CTACTCTCCCTTCTGGGCTCACTATGCTGCGCCAGTGGGACCTTTGGAATACATG 539

DB 241 CTACTCTCCCTTCTGGGCTCACTATGCTGCGCCAGTGGGACCTTTGGAATACATG 300

QY 540 TGTCAACTCTTGACAGGCTCTATTTATAGGCTCTCTCTGGAATCTTCTTCATCATC 599

DB 301 TGTCAACTCTTGACAGGCTCTATTTATAGGCTCTCTCTGGAATCTTCTTCATCATC 360

QY 600 CTCCTGCAATCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659

DB 361 CTCCTGCAATCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

QY 660 AGGCTGACCTTTGGGCTGTGACAGTGTGATCATCTTGGTGGTGGCTGTGCTGCTGCT 719

DB 421 AGGCTGACCTTTGGGCTGTGACAGTGTGATCATCTTGGTGGTGGCTGTGCTGCTGCT 480

QY 720 CTCCTGCAATCATCTTTACAGATCTCAAAAAGAGTCTTCAATACACCTGAGCTCT 779

DB 481 CTCCTGCAATCATCTTTACAGATCTCAAAAAGAGTCTTCAATACACCTGAGCTCT 540

QY 780 CATTTTCAT-----ACATAAAGATGATCATC 807

DB 541 CATTTTCATACAGTCAGTATCAATTTCTGGAAGATTTCCAGACATTAAGATGATCATC 600

QY 808 TTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 867

DB 601 TTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660

QY 868 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 927

DB 661 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720

QY 928 ATGATGTTTATTTCTTCTTCTGGGCTCCCTACAAATGTCCTTCTCTGAAACCTTC 987

DB 721 ATGATGTTTATTTCTTCTTGGGCTCCCTACACATGTCCTTCTCTGACACCTTC 780

QY 988 CAGGAATCTTTGGCCTGAATAATTCAGTAGCTCTAACAGTTGGACCAAGCTATGAG 1047

DB 781 CAGGAATCTTTGGCCTGAATAATTCAGTAGCTCTAACAGTTGGACCAAGCTATGAG 840

QY 1048 GTGACAGAGACTCTTGGGATGACGACTGCTGATCAACCCCAATCACTATGCTTTGTC 1107

DB 841 GTGACAGAGACTCTTGGGATGACGACTGCTGATCAACCCCAATCACTATGCTTTGTC 900

QY 1108 GGGGAGAGTTTCAGAACTACCTCTTACTTCTTCCAAAGACATTCGCAAGGCTTC 1167

DB 901 GGGGAGAGTTTCAGAACTACCTCTTACTTCTTCCAAAGACATTCGCAAGACCTTC 960

QY 1168 TGCAAATCTGTTCTTATTTTCCAGCAAGGCTCCGAGCGAGCAAGCTAGTTTACAC 1227

DB 961 TGCAAATCTGTTCTTATTTTCCAGCAAGGCTCCGAGCGAGCAAGCTAGTTTACACT 1020

QY 1228 CGATCCACTGGGAGCAGGAAATATCTGGGCTTGTGA 1266

DB 1021 CGATCCACTGGGAGCAGGAAATATCTGGGCTTGTGA 1059

RESULT 12

US-08-833-752-1

; Sequence 1, Application US/08833752

; Patent No. 6448375

; GENERAL INFORMATION:

; APPLICANT: SAMSON, MICHEL

; APPLICANT: PARMANTIER, MARC

; APPLICANT: VASSART, GILBERT

; APPLICANT: LIBERT, FREDERICK

; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR

; TITLE OF INVENTION: AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson &amp; Bear

; STREET: 620 Newport Center Drive 16th Floor

; CITY: Newport Beach

; STATE: CA

; COUNTRY: U.S.A.

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/833,752

; FILING DATE: 9-APR-1997

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Altman, Daniel E

; REGISTRATION NUMBER: 34,115

; REFERENCE/DOCKET NUMBER:

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 792 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 240..791

; US-08-833-752-1

Query Match 54.9%; Score 792; DB 4; Length 792;

Best Local Similarity 100.0%; Pred. No. 8.3e-219;

Matches 792; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATCCCAACAGAGCCAAAGCTCTCCATCTAGTGACAGGAGAGCTAGCAGCAAACT 60







QY 977 TGAACACCTCCAGGAATCTTTGGCGCTGATTAATTCAGTAGCTCTTAACAGGTGGACC 1036  
|||||  
Db 874 TGACACCTCCAGGAATCTTTGGCGCTGATTAATTCAGTAGCTCTTAACAGGTGGACC 933  
|||||  
QY 1037 AAGCTATGAGGTGACAGACTTTGGGATGAGCGACTCTGCATCAACCCATCATCT 1096  
|||||  
Db 934 AAGCCAGGAGGTGACAGACTTTGGGATGAGCGACTCTGCATCAATCCATCATCT 993  
|||||  
QY 1097 ATGCTTTGCGGGGAGAGTTCCAGAACTACCTCTTAGTCTTTTCCAAAAGCAGATTG 1156  
|||||  
Db 994 ATGCTTTGCGGGGAGAGTTCCAGAACTATCTCTCGGTCTTCTCCGAAAGCAGATCA 1053  
|||||  
QY 1157 CCAAGAGCTTCGAAATGCTTTCTATTTTCCAGAGAGGCTCCGAGGAGCAGCT 1216  
|||||  
Db 1054 CCAAGAGCTTCGAAATGCTTTCTATTTTCCAGAGAGGCTCCGAGGAGCAGCT 1113  
|||||  
QY 1217 CAGTTTACACCGATCCACTGGGAGCAGAGAAATCTGTGGGCTTTGTACACGGACTCA 1276  
|||||  
Db 1114 CAACAAACAGGCTTCCACTGGGAGCAGAGAGTCTCGGCTGTTTATAAAGCAGGAGC 1173  
|||||  
QY 1277 AGTGGCTGCT 1287  
|||||  
Db 1174 AGTTGATTGT 1184  
|||||

## RESULT 15

US-08-446-669-3  
; Sequence 3, Application US/08446669  
; Patent No. 6132987  
; GENERAL INFORMATION:  
; APPLICANT: Charo, Israel  
; APPLICANT: Coughlin, Shaun  
; TITLE OF INVENTION: MAMMALIAN MONOCYTE CHEMOATTRACTANT  
; TITLE OF INVENTION: PROTEIN RECEPTORS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum  
; STREET: 5 Palo Alto Square  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94306-2155  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,669  
; FILING DATE: May 25, 1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Neeley, Richard  
; REGISTRATION NUMBER: 30,092  
; REFERENCE/DOCKET NUMBER: UCAL-237/01US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-843-5000  
; TELEFAX: 415-857-0663  
; TELEX: 38016CooleyPA  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1979 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 81..1160  
US-08-446-669-3

Query Match 47.4%; Score 683.4; DB 3; Length 1979;  
Best Local Similarity 80.0%; Pred. No. 2.8e-187;  
Matches 825; Conservative 0; Mismatches 186; Indels 20; Gaps 1;  
QY 277 ATTATTTATACATCGGAGCCCTGCCAAAAAATCAATGTAAGCAAAATCGAGCCGCCCTCC 336  
|||||  
Db 154 ATTATGATTAAGGTGCTCCCTGTCATAAATTTGAGCTGAAGCAAAATGGGGGCCAACTCC 213  
|||||  
QY 337 TGCCTCCGCTCTACTCACTGGTGTTCATCTTTGGTGTGTTTGGGCAACATGCTGGTGCATCC 396  
|||||  
Db 214 TGCCTCCGCTCTACTCACTGGTGTTCATCTTTGGTGTGTTTGGGCAACATGCTGGTGCATCC 273  
|||||  
QY 397 TCATCTGATAAATCTGCAAAAGGCTGAAGAGCATGACTGACATCTACCTGCTCAACCTGG 456  
|||||  
Db 274 TCATCTGATAAATCTGCAAAAGGCTGAAGAGCATGACTGACATCTACCTGCTCAACCTGG 333  
|||||  
QY 457 CCATCTGTGACCTGTTTTTTCCTTCTTACTCTCCCTCTGGGCTCACTATGTCGCGGCC 516  
|||||  
Db 334 CCATCTGTGATCTGCTTTTCTTATTAATCTCCATTTGAGGCTCACTCTGCTGCAATG 393  
|||||  
QY 517 AGTGGAGCTTTGGAATACAAATGTGCAACTCTTGACAGGCTCTATTTTATPAGCTTCT 576  
|||||  
Db 394 AGTGGAGCTTTGGAATACAAATGTGCAACTCTTGACAGGCTCTATTTTATPAGCTTCT 453  
|||||  
QY 577 TCTCTGGAATCTTCTTCATCATCTCCTGCTGACATCGATAGTACCTGGCTGCTGCTCATG 636  
|||||  
Db 454 TTGGCGGAATCTTCTTCATCATCTCCTGCTGACATCGATAGTACCTGGCTGCTGCTCATG 513  
|||||  
QY 637 CTGTGTTTGGCTTTAAAGCCAGGAGCTCACCTTTGGGGTGGTGACAAAGTGTGATCATCT 696  
|||||  
Db 514 CTGTGTTTGGCTTTAAAGCCAGGAGCTCACCTTTGGGGTGGTGACAAAGTGTGATCATCT 573  
|||||  
QY 697 GGGTGGTGGCTGTGTTTGGCTCTCCAGGAATCATCTTTACAGATCTCAAAAGAG 756  
|||||  
Db 574 GGTGGTGGCTGTGTTTGGCTCTCCAGGAATCATCTTTACTAAATGCCAGAAAG 633  
|||||  
QY 757 GTCCTCATTAACCTGACGCTCTCAT-----TTTCCATACATTA 796  
|||||  
Db 634 ATTCGTTTATGCTGTGCGCTTATTTCCAGGAGGATGGAATATTTCCACATTA 693  
|||||  
QY 797 AGATAGTCACTTTGGGGCTGGTCTGCCGTGCTTGTGATGCTATCTGCTACTCGGAA 856  
|||||  
Db 694 TGAGGAACATTTGGGGCTGGTCTGCCGTGCTCATCATGCTCATCTGCTACTCGGAA 753  
|||||  
QY 857 TCTTAAATCTGCTTGGTGTGAAATCAGAAAGAGGACAGGCTGTGAGGCTTA 916  
|||||  
Db 754 TCTTAAATCTGCTTGGTGTGAAATCAGAAAGAGGACAGGCTGTGAGGCTTA 813  
|||||  
QY 917 TCTTCAACATCATGATGTTTATTTTCTTCTGCGGCTCCCTACAAATGTCCTTCTCC 976  
|||||  
Db 814 TCTTCAACATCATGATGTTTATTTTCTTCTGCGGCTCCCTACAAATGTCCTTCTCC 873  
|||||  
QY 977 TGAACACCTTCAGGAATCTTTGGCGCTGAATTAATTTGAGTAGTCTTAACAGGTGGACC 1036  
|||||  
Db 874 TGAACACCTTCAGGAATCTTTGGCGCTGAATTAATTTGAGTAGTCTTAACAGGTGGACC 933  
|||||  
QY 1037 AAGCTATGAGGTGACAGACTCTTTGGGATGAGCGACTCTGCATCAACCCATCATCT 1096  
|||||  
Db 934 AAGCCAGGAGGTGACAGACTCTTTGGGATGAGCGACTCTGCATCAATCCATCATCT 993  
|||||  
QY 1097 ATGCTTTGCGGGGAGAGTTCCAGAACTACCTCTTAGTCTTTTCCAAAAGCAGATTG 1156  
|||||  
Db 994 ATGCTTTGCGGGGAGAGTTCCAGAACTATCTCTCGGTCTTCTCCGAAAGCAGATCA 1053  
|||||  
QY 1157 CCAAGAGCTTCGAAATGCTTTCTATTTTCCAGAGAGGCTCCGAGGAGCAGCT 1216  
|||||  
Db 1054 CCAAGAGCTTCGAAATGCTTTCTATTTTCCAGAGAGGCTCCGAGGAGCAGCT 1113  
|||||  
QY 1217 CAGTTTACACCGATCCACTGGGAGCAGAGAAATCTGTGGGCTTTGTACACGGACTCA 1276  
|||||  
Db 1114 CAACAAACAGGCTTCCACTGGGAGCAGAGAGTCTCGGCTGTTTATAAAGCAGGAGC 1173  
|||||

QY 1277 AGTGGCTGGT 1287  
||| | | |  
Db 1174 AGTTGATTGT 1184

Search completed: June 8, 2003, 09:30:46  
Job time : 82 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 08:51:52 ; Search time 233 Seconds  
(without alignments)  
9049.923 Million cell updates/sec

Title: US-09-938-703-3

Perfect score: 1442  
Sequence: 1 GAATGCCCAAGAGGCA.....AGTAGATTAGTCGAATTC 1442

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 870385 seqs, 699768693 residues

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubna/PC1\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubna/PC1US\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/2/pubna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1442	100.0	1442	10	US-09-938-719-3
2	1442	100.0	1442	10	US-09-939-226-3
3	1442	100.0	1442	10	US-09-938-703-3
4	1382.8	95.9	1477	10	US-09-759-841-1
5	1382.8	95.9	1477	10	US-09-938-719-2
6	1382.8	95.9	1477	10	US-09-939-226-2
7	1382.8	95.9	1477	10	US-09-938-703-2
8	1375.2	95.4	143068	9	US-09-967-768A-316
9	1302	90.3	1376	9	US-10-086-814-2
10	1302	90.3	1376	10	US-09-796-202-2
11	1293.4	89.7	1414	9	US-10-232-686-1
12	1293.4	89.7	1414	10	US-09-725-285-1
13	1293.4	89.7	1414	10	US-09-195-662A-1
14	1293.4	89.7	1414	10	US-09-339-912A-1
15	1293.4	89.7	1414	10	US-09-502-783A-1
16	1288.6	89.4	1414	9	US-10-067-800-1
17	1288.6	89.4	1414	10	US-09-779-879A-1
18	1288.6	89.4	1414	10	US-09-779-880A-1
19	1149	79.7	3383	9	US-09-734-221A-13

20 1149 79.7 3383 12 US-10-106-623-1 Sequence 1, Appli  
21 1117 77.5 1225 10 US-09-813-653-14 Sequence 14, Appli  
22 1115.4 77.4 1225 10 US-09-813-653-16 Sequence 16, Appli  
23 980.4 68.0 1056 9 US-10-067-800-21 Sequence 21, Appli  
24 980.4 68.0 1056 10 US-09-779-879A-21 Sequence 21, Appli  
25 980.4 68.0 1056 10 US-09-779-880A-21 Sequence 21, Appli  
26 948.2 65.8 1059 12 US-10-106-623-19 Sequence 19, Appli  
27 792 54.9 792 10 US-09-938-719-1 Sequence 1, Appli  
28 792 54.9 792 10 US-09-938-703-1 Sequence 1, Appli  
29 792 54.9 792 10 US-09-938-703-1 Sequence 1, Appli  
30 680 47.2 1083 10 US-09-131-827A-19 Sequence 19, Appli  
31 678.4 47.0 1083 10 US-09-131-827A-1 Sequence 1, Appli  
32 320.2 22.1 1689 10 US-09-931-361A-15 Sequence 15, Appli  
33 319.4 22.1 1717 10 US-09-964-824A-100 Sequence 100, App  
34 318.2 22.1 1915 12 US-10-106-623-3 Sequence 3, Appli  
35 318.2 22.1 1065 9 US-09-922-895-2 Sequence 2, Appli  
36 307.4 21.3 3426 9 US-10-001-835-29 Sequence 29, Appli  
37 294.8 20.4 1607 9 US-10-120-394-19 Sequence 19, Appli  
38 294.8 20.4 1607 9 US-09-764-413-19 Sequence 19, Appli  
39 294.8 20.4 1677 10 US-09-837-446-1 Sequence 1, Appli  
40 237.2 16.4 1318 10 US-09-917-800A-1445 Sequence 1445, Ap  
41 229.8 15.9 1586 10 US-09-104-792-1 Sequence 1, Appli  
42 194.6 13.5 1487 10 US-09-789-482-3 Sequence 3, Appli  
43 194.6 13.5 1487 10 US-09-789-486-3 Sequence 3, Appli  
44 194.6 13.5 3100 10 US-09-954-456-267 Sequence 267, App  
45 194.6 13.5 3100 10 US-09-954-456-945 Sequence 945, App

#### ALIGNMENTS

#### RESULT 1

US-09-938-719-3  
; Sequence 3, Application US/09938719  
; Patent No. US20020106742A1  
; GENERAL INFORMATION:  
; APPLICANT: SAMSON, MICHEL  
; PARMENTIER, MARC  
; VASSART, GILBERT  
; LIBERT, FREDERICK

TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR

NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 620 Newport Center Drive 16th Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 92660

#### COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/938,719

FILING DATE: 24-Aug-2001  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/626,939

FILING DATE: 27-JULY-2000  
ATTORNEY/AGENT INFORMATION:

NAME: Altman, Daniel E  
REGISTRATION NUMBER: 34,115

REFERENCE/DOCKET NUMBER: <Unknown>

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1442 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)



INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1442 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 240..884  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-939-226-3

Query Match 100.0%; Score 1442; DB 10; Length 1442;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1442; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAATCCCCCAACAGAGCCAGCTCCATCTAGTGACAGGGAAGCTAGCAGCAACCT 60  
DB 1 GAATCCCCCAACAGAGCCAGCTCCATCTAGTGACAGGGAAGCTAGCAGCAACCT 60  
QY 61 TCCCTTCACATACAAAACCTCATTGCTGGCCAAAAGAGAGTTAATCAATGTAGCATC 120  
DB 61 TCCCTTCACATACAAAACCTCATTGCTGGCCAAAAGAGAGTTAATCAATGTAGCATC 120  
QY 121 TATGTAGCAATTAACCACTTATGATGTATATAACAGAGTTTGCATTCATGAGGGCACT 180  
DB 121 TATGTAGCAATTAACCACTTATGATGTATATAACAGAGTTTGCATTCATGAGGGCACT 180  
QY 181 AAATACATCTAGGACTTATAAAGATCAGCTTTTATATATGACAGAGGTGGAAAGA 240  
DB 181 AAATACATCTAGGACTTATAAAGATCAGCTTTTATATATGACAGAGGTGGAAAGA 240  
QY 241 TGGATTATCAAGTGTCAAGTCCAACTATGACATCAATATTATATACATGGAGCCCTGCC 300  
DB 241 TGGATTATCAAGTGTCAAGTCCAACTATGACATCAATATTATATACATGGAGCCCTGCC 300  
QY 301 AAAAATCAATGTGAAGCAAAATCGCAGCCCGCTCGCTCGCTCTACTCACTGGTGT 360  
DB 301 AAAAATCAATGTGAAGCAAAATCGCAGCCCGCTCGCTCGCTCTACTCACTGGTGT 360  
QY 361 TCATCTTTGGTTTGGGGCAACATGCTGGTCACTCCCTGATCAATCACTGATAAAGGC 420  
DB 361 TCATCTTTGGTTTGGGGCAACATGCTGGTCACTCCCTGATCAATCACTGATAAAGGC 420  
QY 421 TGAAGAGCATGACATGACATCTACCTGCTCAACCTGGCCATCTGACCTGTTTTCCTC 480  
DB 421 TGAAGAGCATGACATGACATCTACCTGCTCAACCTGGCCATCTGACCTGTTTTCCTC 480  
QY 481 TTACTGTCCCTTCTGGGCTCAGTATGCTGCGCCGAGTGGAGCTTTGGAATACAATGT 540  
DB 481 TTACTGTCCCTTCTGGGCTCAGTATGCTGCGCCGAGTGGAGCTTTGGAATACAATGT 540  
QY 541 GTCAACTCTTGACAGGCTCTATTATATAGGCTTCTCTGGAATCTTCTCATCATCC 600  
DB 541 GTCAACTCTTGACAGGCTCTATTATATAGGCTTCTCTGGAATCTTCTCATCATCC 600  
QY 601 TCCGACATPCGATAGGTACCTGCTGTCTGCTCCATGCTGTGTTTAAAGCCAGGA 660  
DB 601 TCCGACATPCGATAGGTACCTGCTGTCTGCTCCATGCTGTGTTTAAAGCCAGGA 660  
QY 661 CGGTACCTTTGGGTGGTGACAGTGTGATCAGTGGGTGGTGGCTGTGTTGGCTCTC 720  
DB 661 CGGTACCTTTGGGTGGTGACAGTGTGATCAGTGGGTGGTGGCTGTGTTGGCTCTC 720  
QY 721 TCCAGAGATCACTTTACAGATCTCAAAAGAGGTCTTCATTCACCTGCGAGCTCTC 780  
DB 721 TCCAGAGATCACTTTACAGATCTCAAAAGAGGTCTTCATTCACCTGCGAGCTCTC 780  
QY 781 ATTTTCCATACATTAAGATAGTCACTTTGGGGCTGTGCTGCGCTGCTGCTCATGCTC 840  
DB 781 ATTTTCCATACATTAAGATAGTCACTTTGGGGCTGTGCTGCGCTGCTGCTCATGCTC 840

QY 841 ATCTGCTACTCGGGAATCTAAAACTCTGCTCGTCTCGAATGAGAAAGAGGCAC 900  
DB 841 ATCTGCTACTCGGGAATCTAAAACTCTGCTCGTCTCGAATGAGAAAGAGGCAC 900  
QY 901 AGGCTGTGAGGCTTATCTTCCATCATGATGTTTATTTCTCTCTGCGGCTCCCTAC 960  
DB 901 AGGCTGTGAGGCTTATCTTCCATCATGATGTTTATTTCTCTCTGCGGCTCCCTAC 960  
QY 961 AACATGCTCTCTCTGAGACACCTTCCAGGATCTTTGGCTGATATATGAGTAGC 1020  
DB 961 AACATGCTCTCTCTGAGACACCTTCCAGGATCTTTGGCTGATATATGAGTAGC 1020  
QY 1021 TCTAACAGGTTGGACCAAGCTATGACGTCACAGAGACTCTTGGATGACGACTGTGC 1080  
DB 1021 TCTAACAGGTTGGACCAAGCTATGACGTCACAGAGACTCTTGGATGACGACTGTGC 1080  
QY 1081 ATCAACCCCATCATCTATGCTTGTGCGGAGAGTTCAGAACTACCTCTAGCTTC 1140  
DB 1081 ATCAACCCCATCATCTATGCTTGTGCGGAGAGTTCAGAACTACCTCTAGCTTC 1140  
QY 1141 TTCCAAAAGACACATTCGCAACGCTTCTGCAATGCTTCTATTTTCCAGCAAGGCT 1200  
DB 1141 TTCCAAAAGACACATTCGCAACGCTTCTGCAATGCTTCTATTTTCCAGCAAGGCT 1200  
QY 1201 CCCGAGCGAGCAAGCTCAGTTTACACCCGATCCACTGGGAGCAGGAAATATCTG 1260  
DB 1201 CCCGAGCGAGCAAGCTCAGTTTACACCCGATCCACTGGGAGCAGGAAATATCTG 1260  
QY 1261 TTGTACACGAGCTCAAGTGGGCTGGTGGCTGGTGGCTGGTGGCTGGCTGGCT 1320  
DB 1261 TTGTACACGAGCTCAAGTGGGCTGGTGGCTGGTGGCTGGTGGCTGGCTGGCT 1320  
QY 1321 TCATACACAGCTGGGCTGGGCTGGTGGGAGTCTTTTAAAGGAAGTACTGTTA 1380  
DB 1321 TCATACACAGCTGGGCTGGGCTGGGAGTCTTTTAAAGGAAGTACTGTTA 1380  
QY 1381 TAGAGGCTTAAGATTCATCCATTTATTTGGCATCTGTTTAAAGTAGATAGCCGAAT 1440  
DB 1381 TAGAGGCTTAAGATTCATCCATTTATTTGGCATCTGTTTAAAGTAGATAGCCGAAT 1440  
QY 1441 TC 1442  
DB 1441 TC 1442  
RESULT 3  
US-09-938-703-3  
Sequence 3, Application US/09938703  
Patent No. US20020110870A1  
GENERAL INFORMATION:  
APPLICANT: SAMSON, MICHEL  
PARMENTIER, MARC  
VASSART, GILBERT  
LIBERT, FREDERICK  
TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR  
AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 620 Newport Center Drive 16th Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/938,703  
FILING DATE: 24-Aug-2001  
CLASSIFICATION: <Unknown>

```

P R I O R   A P P L I C A T I O N   D A T A :
A P P L I C A T I O N   N U M B E R : 09/626,939
F I L I N G   D A T E : 2000-07-27
A T T O R N E Y / A G E N T   I N F O R M A T I O N :
N A M E : A l t m a n , D a n i e l   E
R E G I S T R A T I O N   N U M B E R : 34,115
R E F E R E N C E / D O C K E T   N U M B E R : <Unknown>
I N F O R M A T I O N   F O R   S E Q   I D   N O : 3 :
S E Q U E N C E   C H A R A C T E R I S T I C S :
L E N G T H : 1442   b a s e   p a i r s
T Y P E : n u c l e i c   a c i d
S T R A N D E D N E S S : s i n g l e
T O P O L O G Y : l i n e a r
M O L E C U L E   T Y P E : D N A   ( g e n o m i c )
F E A T U R E :
N A M E / K E Y : C D S
L O C A T I O N : 240..884
S E Q U E N C E   D E S C R I P T I O N : S E Q   I D   N O : 3 :
U S - 09 - 938 - 703 - 3

```

```

RESULT 4
US-09-759-841-1
? Sequence 1, Application US/09759841
? Patent No. US20010039206A1
? GENERAL INFORMATION:
? APPLICANT: Bickett, Graham A
? APPLICANT: Dobbs, Susan
? APPLICANT: Perros, Manoussos
? TITLE OF INVENTION: Assay Method
? FILE REFERENCE: PC10349APME
? CURRENT APPLICATION NUMBER: US/09759,841
? CURRENT FILING DATE: 2001-01-12
? PRIOR APPLICATION NUMBER: GB 0000661.9
? PRIOR FILING DATE: 2000-01-12
? PRIOR APPLICATION NUMBER: GB 0000663.5
? PRIOR FILING DATE: 2000-01-12
? PRIOR APPLICATION NUMBER: GB 0000659.3
? PRIOR FILING DATE: 2000-01-12
? NUMBER OF SEQ IDS: 6
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 1

```



ATTORNEY/AGENT INFORMATION:

NAME: Altman, Daniel E  
REGISTRATION NUMBER: 34,115  
REFERENCE/DOCKET NUMBER: <Unknown>  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1477 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 240..1295  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-938-719-2

Query Match 95.9%; Score 1382.8; DB 10; Length 1477;  
Best Local Similarity 97.4%; Pred. No. 0;  
Matches 1439; Conservative 0; Mismatches 3; Indels 35; Gaps 2;

QY 1 GAATTCCTCCACAGAGCAAGCTCTCCATCTAGTGGACAGGAGCTAGCAGCAACCT 60  
DB 1 GAATTCCTCCACAGAGCAAGCTCTCCATCTAGTGGACAGGAGCTAGCAGCAACCT 60  
QY 61 TCCCTTCACACAACTCAATGCTTGGCCAAAAGAGAGTTAATCAATGTAGACATC 120  
DB 61 TCCCTTCACACAACTCAATGCTTGGCCAAAAGAGAGTTAATCAATGTAGACATC 120  
QY 121 TATGTAGGCAATTAACCACTATTGATGATATAACAGTTTGCATTCATGGAGGCACT 180  
DB 121 TATGTAGGCAATTAACCACTATTGATGATATAACAGTTTGCATTCATGGAGGCACT 180  
QY 181 AAATACATCTAGGACTTTAAGAGATCAGCTTTTATATGACAGAGGTGGAAAGA 240  
DB 181 AAATACATCTAGGACTTTAAGAGATCAGCTTTTATATGACAGAGGTGGAAAGA 240  
QY 241 TGGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATATATATACATCGAGCCCTGCC 300  
DB 241 TGGATTATCAAGTGTCAAGTCCCAATCTATGACATCAATATATATACATCGAGCCCTGCC 300  
QY 301 AAAAATCAATGTGAAGCAATCGACGCCCTCTCGCTCGGCTCTACTCACTGGTGT 360  
DB 301 AAAAATCAATGTGAAGCAATCGACGCCCTCTCGCTCGGCTCTACTCACTGGTGT 360  
QY 361 TCATCTTGGTTTGTGGGCAACATGCTGTCATCTCACTCACTGATTAATACAGTCAAGGC 420  
DB 361 TCATCTTGGTTTGTGGGCAACATGCTGTCATCTCACTCACTGATTAATACAGTCAAGGC 420  
QY 421 TGAAGAGCATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTTC 480  
DB 421 TGAAGAGCATGACTGACATCTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTCCTTC 480  
QY 481 TTACTGTCCCTTCTGGGCTCACTGCTGCGGCCAGTGGGACTTTGGAATACATGTT 540  
DB 481 TTACTGTCCCTTCTGGGCTCACTGCTGCGGCCAGTGGGACTTTGGAATACATGTT 540  
QY 541 GTCACTCTTGACAGGCTCTATTTATAGGCTTCTCTCGGAATCTTCTTCATCATCC 600  
DB 541 GTCACTCTTGACAGGCTCTATTTATAGGCTTCTCTCGGAATCTTCTTCATCATCC 600  
QY 601 TCGTGAATCGATAGTACCTGGCTGCTGCTCATGCTGTTTGTGTTTAAAGCCAGGA 660  
DB 601 TCGTGAATCGATAGTACCTGGCTGCTGCTCATGCTGTTTGTGTTTAAAGCCAGGA 660  
QY 661 CGGTCACCTTTGGGTTGGTGAAGTGTATCACTTGGGTTGGTGGCTGTTGGGCTC 720  
DB 661 CGGTCACCTTTGGGTTGGTGAAGTGTATCACTTGGGTTGGTGGCTGTTGGGCTC 720  
QY 721 TCCAGGAATCATCTTTACAGATCTCAAAAAGAGGCTTCATTAACCTGCAGCTCTC 780  
DB 721 TCCAGGAATCATCTTTACAGATCTCAAAAAGAGGCTTCATTAACCTGCAGCTCTC 780  
QY 781 ATTTCAT-----ACATTAAGATAGTCACT 808

DB 781 ATTTCATACAGTCAGTATCAATTTCTGGAAGATTTCCAGACATTAAGATAGTCACT 840  
QY 809 TGGGCTGGTCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 868  
DB 841 TGGGCTGGTCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
QY 869 TGTTCGGTGTGCGAATGAGAAGAGAGGACAGGCTGTGAGGCTTATCTTCACCATCA 928  
DB 901 TGTTCGGTGTGCGAATGAGAAGAGAGGACAGGCTGTGAGGCTTATCTTCACCATCA 960  
QY 929 TGATTTGTTTATTTCTTCTTGGGCTCCCTTACAACTTCTCTTCTTCTTCTTCTTCTT 988  
DB 961 TGATTTGTTTATTTCTTCTTGGGCTCCCTTACAACTTCTCTTCTTCTTCTTCTTCTT 1020  
QY 989 AGGAATCTTTGGGCTGAATATTTGAGTGTGCTTACAGGTTGGACCAAGTATGCAAG 1048  
DB 1021 AGGAATCTTTGGGCTGAATATTTGAGTGTGCTTACAGGTTGGACCAAGTATGCAAG 1080  
QY 1049 TGACAGAGACTCTTGGGATGAGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1108  
DB 1081 TGACAGAGACTCTTGGGATGAGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
QY 1109 GGGAGAGTTGAGAACTGCTTGTAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1168  
DB 1141 GGGAGAGTTGAGAACTGCTTGTAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1200  
QY 1169 GCAATCTGCTTCTTATTTCCAGAGAGGCTCCGAGGAGAGGCTGAGTTCACACCC 1228  
DB 1201 GCAATCTGCTTCTTATTTCCAGAGAGGCTCCGAGGAGAGGCTGAGTTCACACCC 1260  
QY 1229 GATCCACTGGGAGAGAGAAATATCTGTTGGGCTGTGACACGAGTCAAGTGGGCTGGT 1288  
DB 1261 GATCCACTGGGAGAGAGAAATATCTGTTGGGCTGTGACACGAGTCAAGTGGGCTGGT 1320  
QY 1289 ACCAGTCAAGAGTTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1348  
DB 1321 ACCAGTCAAGAGTTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380  
QY 1349 GG---GAGTCTTTTAAAGAGAGTACTGTTATAGAGGCTTAAGATTCATCCATTT 1405  
DB 1381 TGGNAGAGTCTTTTAAAGAGAGTACTGTTATAGAGGCTTAAGATTCATCCATTT 1440  
QY 1406 ATTGGCATCTGTTTAAAGAGAGTACTGTTATAGAGGCTTAAGATTCATCCATTT 1442  
DB 1441 ATTGGCATCTGTTTAAAGAGAGTACTGTTATAGAGGCTTAAGATTCATCCATTT 1477

## RESULT 6

US-09-939-226-2  
Sequence 2, Application US/09939226  
Patent No. US20020110805A1  
GENERAL INFORMATION:  
APPLICANT: SAMSON, MICHEL  
VASSART, GILBERT  
LIBERT, FREDERICK  
TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR  
AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 620 Newport Center Drive 16th Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: U.S.A.  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/939,226  
FILING DATE: 24-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIORITY DATA:  
APPLICATION NUMBER: 09/626,939  
FILING DATE: 2000-07-27  
ATTORNEY/AGENT INFORMATION:  
NAME: Altman, Daniel E  
REGISTRATION NUMBER: 34,115  
REFERENCE/DOCKET NUMBER: <Unknown>  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1477 base pairs  
TYPE: nucleic acid  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 240..1295  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-939-226-2

Query Match 95.98; Score 1382.8; DB 10; Length 1477;  
Best Local Similarity 97.48; Pred. No. 0;  
Matches 1439; Conservative 0; Mismatches 3; Indels 35; Gaps 2;

QY 1 GAATCCCCCAGAGCCAGCTCTCCATCTAGTGGACAGGAAAGCTAGCAGCAAACT 60  
DB 1 GAATCCCCCAGAGCCAGCTCTCCATCTAGTGGACAGGAAAGCTAGCAGCAAACT 60  
QY 61 TCCCTTCACTACAACTTCACTGCTTGGCCAAAAGAGAGTAACTCATGTAGACATC 120  
DB 61 TCCCTTCACTACAACTTCACTGCTTGGCCAAAAGAGAGTAACTCATGTAGACATC 120  
QY 121 TATGTAGGCAATTAACCACTATTGATGATATAAAGAGAGTTCATCATGAGGCAACT 180  
DB 121 TATGTAGGCAATTAACCACTATTGATGATATAAAGAGAGTTCATCATGAGGCAACT 180  
QY 181 AATATATCTAGAGCTTTAATAAGATCACTTTTATTTATGACAGGGTGGAAACAAG 240  
DB 181 AATATATCTAGAGCTTTAATAAGATCACTTTTATTTATGACAGGGTGGAAACAAG 240  
QY 241 TGATATACAGTGTCAAGTCAATCTATGACATCAATATTATATATATGAGCCCTGCC 300  
DB 241 TGATATACAGTGTCAAGTCAATCTATGACATCAATATTATATATATGAGCCCTGCC 300  
QY 301 AAAAAATCAATGTGAAGCAATCGCAGCCCGCTCTGCTCCGCTCTACTCTACTGGTGT 360  
DB 301 AAAAAATCAATGTGAAGCAATCGCAGCCCGCTCTGCTCCGCTCTACTCTACTGGTGT 360  
QY 361 TCATCTTTGGTTTGGGCAACATGCTGGTCACTCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
DB 361 TCATCTTTGGTTTGGGCAACATGCTGGTCACTCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
QY 421 TGAAGAGATGACTGACATCTACTGCTCAACCTGGCCATCTGACCTGTTTCTCTC 480  
DB 421 TGAAGAGATGACTGACATCTACTGCTCAACCTGGCCATCTGACCTGTTTCTCTC 480  
QY 481 TTACTGTCCCTCTGGGCTCACTATGCTGCGCCAGTGGGACCTTTGGAATACAACT 540  
DB 481 TTACTGTCCCTCTGGGCTCACTATGCTGCGCCAGTGGGACCTTTGGAATACAACT 540  
QY 541 GTCACTCTGACAGGCTCTATTTATAGGCTTCTCTGGAATCTTCTCATCATCC 600  
DB 541 GTCACTCTGACAGGCTCTATTTATAGGCTTCTCTGGAATCTTCTCATCATCC 600  
QY 601 TCTGACATGATAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
DB 601 TCTGACATGATAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660  
QY 661 CGGTCACTTTGGGGTGGTGACAAAGTGTATGATCACTGGGTGGTGGTGGTGGTGGTGGT 720  
DB 661 CGGTCACTTTGGGGTGGTGACAAAGTGTATGATCACTGGGTGGTGGTGGTGGTGGTGGT 720

DB 661 CGGTCACTTTGGGGTGGTGACAAAGTGTATGATCACTGGGTGGTGGTGGTGGTGGTGGT 720  
QY 721 TCCAGGAATCATCTTTACAGATCTCAAAAAGAGGCTCTCAATTAACCTGCGAGCTCTC 780  
DB 721 TCCAGGAATCATCTTTACAGATCTCAAAAAGAGGCTCTCAATTAACCTGCGAGCTCTC 780  
QY 781 ATTTTCCAT-----ACATTAAGATAGTATCT 808  
DB 781 ATTTTCCATAGCTCAGTATCAATTTCTGGAAGAAATTTCCAGACATTAAGATAGTATCT 840  
QY 809 TGGGGCTGGTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 868  
DB 841 TGGGGCTGGTCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
QY 869 TGCCTTGGTCTGAAATGAGAAGAGAGCAGAGGCTGTGAGGCTTATCTTCCACATCA 928  
DB 901 TGCCTTGGTCTGAAATGAGAAGAGAGCAGAGGCTGTGAGGCTTATCTTCCACATCA 960  
QY 929 TGATTTTATTTTCTTCTTGGGCTCCCTCAACANTGCTTCTTCTTCCAGACCTTCC 988  
DB 961 TGATTTTATTTTCTTCTTGGGCTCCCTCAACANTGCTTCTTCTTCCAGACCTTCC 1020  
QY 989 AGGAATTTTGGCTGAATAATTTGAGTAGCTCTAAGAGTGGACCAAGCTATGAGG 1048  
DB 1021 AGGAATTTTGGCTGAATAATTTGAGTAGCTCTAAGAGTGGACCAAGCTATGAGG 1080  
QY 1049 TGACAGAGACTCTTGGGAGTGCAGTCTGCTCAACCCCATCATCTATGCTTCTGCTG 1108  
DB 1081 TGACAGAGACTCTTGGGAGTGCAGTCTGCTCAACCCCATCATCTATGCTTCTGCTG 1140  
QY 1109 GGGAGAGTTTCAAACTACCTCTTAGTCTTCTTCCAAAAGACATTCGCAACGCTTCT 1168  
DB 1141 GGGAGAGTTTCAAACTACCTCTTAGTCTTCTTCCAAAAGACATTCGCAACGCTTCT 1200  
QY 1169 GCAATCTCTTCTTATTTTCCAGCAAGAGCTCCCGAGGAGCAAGCTCAGTTTACACCC 1228  
DB 1201 GCAATCTCTTCTTATTTTCCAGCAAGAGCTCCCGAGGAGCAAGCTCAGTTTACACCC 1260  
QY 1229 GATCCACTGGGAGCAGGAATAATCTGTGGGCTGTGACAGGACTCAAGTGGGCTGGT 1288  
DB 1261 GATCCACTGGGAGCAGGAATAATCTGTGGGCTGTGACAGGACTCAAGTGGGCTGGT 1320  
QY 1289 ACCAGTCAAGTGTGACATGCTTAGTCTTATACAGAGCTGGGCTGGGCTGGGCTGGT 1348  
DB 1321 ACCAGTCAAGTGTGACATGCTTAGTCTTATACAGAGCTGGGCTGGGCTGGGCTGGT 1380  
QY 1349 GGGCTCTTTTAAAAGAGAGTACTGCTTATAGAGGCTAAGATTCATCCATTT 1405  
DB 1381 TGGNAGAGTCTTTTAAAAGAGAGTACTGCTTATAGAGGCTAAGATTCATCCATTT 1440  
QY 1405 ATTTGGCACTGTTTAAAGTAGATTAGATCCGAATTC 1442  
DB 1441 ATTTGGCACTGTTTAAAGTAGATTAGATCCGAATTC 1477

RESULT 7

US-09-938-703-2  
; Sequence 2, Application US/09938703  
; Patent No. US20020110870A1  
; GENERAL INFORMATION:  
; APPLICANT: SANSON, MICHEL  
; VASSART, GILBERT  
; LIBERT, FREDERICK  
; TITLE OF INVENTION: ACTIVE AND INACTIVE CC-CHEMOKINES RECEPTOR  
; AND NUCLEIC ACID MOLECULES ENCODING SAID RECEPTOR  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 620 Newport Center Drive 16th Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: U.S.A.



; PRIOR APPLICATION NUMBER: US/60/236,109  
 ; PRIOR FILING DATE: 2000-09-28  
 ; PRIOR APPLICATION NUMBER: US/60/236,034  
 ; PRIOR FILING DATE: 2000-09-28  
 ; PRIOR APPLICATION NUMBER: US/60/236,111  
 ; PRIOR FILING DATE: 2000-09-28  
 ; NUMBER OF SEQ ID NOS: 325  
 ; SOFTWARE: Patent in version 3.0  
 ; SEQ ID NO 316  
 ; LENGTH: 143068  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-967-768A-316

Query Match 95.4%; Score 1375.2; DB 10; Length 143068;  
 Best Local Similarity 97.4%; Pred. No. 0;  
 Matches 1432; Conservative 0; Mismatches 3; Indels 35; Gaps 2;  
 QY 1 GAATTCCTCCCAACAGCCAGCTCTCCATCTAGTGGACAGGAGCTAGCAGCAACT 60  
 Db 61244 GAATTCCTCCCAACAGCCAGCTCTCCATCTAGTGGACAGGAGCTAGCAGCAACT 61303  
 QY 61 TCCTTCACTACAAAATTCATTGCTTGGCCAAAAGAGAGTAATTCATGTAGACATC 120  
 Db 61304 TCCTTCACTACAAAATTCATTGCTTGGCCAAAAGAGAGTAATTCATGTAGACATC 61363  
 QY 121 TATGTAGCAATTAACCACTATTGATGTATAAAGCAGTTTGCATTCATGAGGCAACT 180  
 Db 61364 TATGTAGCAATTAACCACTATTGATGTATAAAGCAGTTTGCATTCATGAGGCAACT 61423  
 QY 181 AATATCATCTAGGACTTTATAAAGATCACTTTTATTATGACAGGTTGGAACAAGA 240  
 Db 61424 AATATCATCTAGGACTTTATAAAGATCACTTTTATTATGACAGGTTGGAACAAGA 61483  
 QY 241 TGGATTATCAAGTGTCAAGTCCAAATCTATGACATCAATTAATTAACATCGGAGCCCTGCC 300  
 Db 61484 TGGATTATCAAGTGTCAAGTCCAAATCTATGACATCAATTAATTAACATCGGAGCCCTGCC 61543  
 QY 301 AAAAATCAATGTAAGCAATTCGACGCCCTCTCCCTCCCTCTACTCTACTCTACTCTCT 360  
 Db 61544 AAAAATCAATGTAAGCAATTCGACGCCCTCTCCCTCCCTCTACTCTACTCTACTCTCT 61603  
 QY 361 TCATCTTTGGTTTGTGGGCAACATGCTGTCTCATCTCTCATCTGATTAATTCGAAAGGC 420  
 Db 61604 TCATCTTTGGTTTGTGGGCAACATGCTGTCTCATCTCTCATCTGATTAATTCGAAAGGC 61663  
 QY 421 TGAAGCATGACATGACATCTACCTGCTCAACCTGGGCAATCTGACCTGTTTTCCTCTC 480  
 Db 61664 TGAAGCATGACATGACATCTACCTGCTCAACCTGGGCAATCTGACCTGTTTTCCTCTC 61723  
 QY 481 TTACTGTCCTCTCTGGGCTCACTATGCTGCGGCCAGTGGGACTTTGGAATACAAATGT 540  
 Db 61724 TTACTGTCCTCTCTGGGCTCACTATGCTGCGGCCAGTGGGACTTTGGAATACAAATGT 61783  
 QY 541 GTCACTCTTGACAGGCTCTATTTATAGGCTTCTCTGGAATCTTCTTCATCATCC 600  
 Db 61784 GTCACTCTTGACAGGCTCTATTTATAGGCTTCTCTGGAATCTTCTTCATCATCC 61843  
 QY 601 TCTGACATCGATAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 560  
 Db 61844 TCTGACATCGATAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 61903  
 QY 661 CGGTCACTTTGGGCTGGTGACAAAGTGTATCACTTGGGTGGTGGTGGTGGTGGTGGTGGT 720  
 Db 61904 CGGTCACTTTGGGCTGGTGACAAAGTGTATCACTTGGGTGGTGGTGGTGGTGGTGGTGGT 61963  
 QY 721 TCCAGGATCATCTTTACAGATCTCAAAAGAGTCTTCATTCACCTGCACTGCTC 780  
 Db 61964 TCCAGGATCATCTTTACAGATCTCAAAAGAGTCTTCATTCACCTGCACTGCTC 62023  
 QY 781 ATTTTCCAT-----ACATTAAGATAGTCACT 808  
 Db 62024 ATTTTCCATAGTATCAATTCGGAAGAAATTCACAGACATTAAGATAGTCACT 62083

QY 809 TGGGGCTGGTCTGCCGCTGCTGTCTCATGGTCACTGCTACTCGGAAATCTAAAAATC 868  
 Db 62084 TGGGGCTGGTCTGCCGCTGCTGTCTCATGGTCACTGCTACTCGGAAATCTAAAAATC 62143  
 QY 869 TGGCTGGGTGCGAATGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 928  
 Db 62144 TGGCTGGGTGCGAATGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 62203  
 QY 929 TGATTTGTTTATTTTCTCTTCTGGGCTCCCTACAACTTGTCTTCTTCTTCTTCTTCTTCT 988  
 Db 62204 TGATTTGTTTATTTTCTTCTTCTGGGCTCCCTACAACTTGTCTTCTTCTTCTTCTTCT 62263  
 QY 989 AGGAATTCCTTGGCTGAAATTAATTCAGTGTCTTAAAGAGGAGGAGGAGGAGGAGGAGGAG 1048  
 Db 62264 AGGAATTCCTTGGCTGAAATTAATTCAGTGTCTTAAAGAGGAGGAGGAGGAGGAGGAGGAG 62323  
 QY 1049 TGACAGAGACTCTTGGGATGACGACTGCTGATCAACCCCACTATCTATGCTTTGTGG 1108  
 Db 62324 TGACAGAGACTCTTGGGATGACGACTGCTGATCAACCCCACTATCTATGCTTTGTGG 62383  
 QY 1109 GGGAGAAGTTCAGAACTTACCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1168  
 Db 62384 GGGAGAAGTTCAGAACTTACCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 62443  
 QY 1169 GCAATGCTGTCTATTTTCCAGAGAGGCTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1228  
 Db 62444 GCAATGCTGTCTATTTTCCAGAGAGGCTTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 62503  
 QY 1229 GATCCACTGGGAGCAGGAGAAATATCTGTGGGCTTGTGACAGGACTCAAGTGGGCTGGTG 1288  
 Db 62504 GATCCACTGGGAGCAGGAGAAATATCTGTGGGCTTGTGACAGGACTCAAGTGGGCTGGTG 62563  
 QY 1289 ACCAGTCAAGTGTGACATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1346  
 Db 62564 ACCAGTCAAGTGTGACATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 62623  
 QY 1347 -TTGGGAGGCTTTTAAAGAGGAGTACTGTTATAGAGGCTTAAAGATTCATCCATTT 1405  
 Db 62624 TGGGAGAGGCTTTTAAAGAGGAGTACTGTTATAGAGGCTTAAAGATTCATCCATTT 62683  
 QY 1406 ATTTGGCATCTGTTTAAAGTAGATTAGATC 1435  
 Db 62684 ATTTGGCATCTGTTTAAAGTAGATTAGATC 62713

RESULT 9  
 US-10-086-814-2  
 ; Sequence 2, Application US/10086814  
 ; Publication No. US20030092632A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dragic, Tatjana  
 ; APPLICANT: Olson, William C.  
 ; TITLE OF INVENTION: SULFATED CCR5 PEPTIDES FOR HIV-1 INFECTION  
 ; FILE REFERENCE: 61010-AB-1  
 ; CURRENT APPLICATION NUMBER: US/10/086,814  
 ; CURRENT FILING DATE: 2002-02-28  
 ; NUMBER OF SEQ ID NOS: 38  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 2  
 ; LENGTH: 1376  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-086-814-2

Query Match 90.3%; Score 1302; DB 9; Length 1376;  
 Best Local Similarity 97.7%; Pred. No. 0;  
 Matches 1344; Conservative 0; Mismatches 0; Indels 32; Gaps 1;  
 QY 1 GAATTCCTCCCAACAGCCAGCTCTCCATCTAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60  
 Db 1 GAATTCCTCCCAACAGCCAGCTCTCCATCTAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60

QY	61	TCCTCTCACTACAAACCTTCATTGCTTGGCCAAAGAGAGTTAATTCATGTGACATC	120
DB	61	TCCTCTCACTACAAACCTTCATTGCTTGGCCAAAGAGAGTTAATTCATGTGACATC	120
QY	121	TATGTAGGCAATTAATAACCTTATTGATGTATATAACAGTTTCATTTCATCGAGGCAACT	180
DB	121	TATGTAGGCAATTAATAACCTTATTGATGTATATAACAGTTTCATTTCATCGAGGCAACT	180
QY	181	AAATACATCTAGGACTTTTAAAGATCACTTTTATTTATGACAGGGTGGAAACAAGA	240
DB	181	AAATACATCTAGGACTTTTAAAGATCACTTTTATTTATGACAGGGTGGAAACAAGA	240
QY	241	TGGATTATCAAGTGTCAAGTCCCACTATATGACATCAATATTATATACATCGAGGCGCTGCC	300
DB	241	TGGATTATCAAGTGTCAAGTCCCACTATATGACATCAATATTATATACATCGAGGCGCTGCC	300
QY	301	AAAAAATCAATGTGAAGCAAAATCGACGCGCGCTCTGGCTCCGCTCTACTCACTGGTGT	360
DB	301	AAAAAATCAATGTGAAGCAAAATCGACGCGCGCTCTGGCTCCGCTCTACTCACTGGTGT	360
QY	361	TCATCTTTGGTTTTGTGGGCAACAATGCTGGTCATCTCATCTGATAAACTGCAAAAGGC	420
DB	361	TCATCTTTGGTTTTGTGGGCAACAATGCTGGTCATCTCATCTGATAAACTGCAAAAGGC	420
QY	421	TGAAGAGCATGACTGACATCTPACTGCTCTCAACCTGGCCATCTCTGACCTGTTTTCCCTTC	480
DB	421	TGAAGAGCATGACTGACATCTPACTGCTCTCAACCTGGCCATCTCTGACCTGTTTTCCCTTC	480
QY	481	TTACTGTCCCTCTCTGGGCTCACTATGCTCGGCGCCAGTGGGACTTTGGAAATACAAATGT	540
DB	481	TTACTGTCCCTCTCTGGGCTCACTATGCTCGGCGCCAGTGGGACTTTGGAAATACAAATGT	540
QY	541	GTCAACTCTTGACAGGGCTCTATTTTATATAGGCTCTCTCTCTGGAATCTTCTTCATCATCC	600
DB	541	GTCAACTCTTGACAGGGCTCTATTTTATATAGGCTCTCTCTCTGGAATCTTCTTCATCATCC	600
QY	601	TCCTGACAATCGATAGGTACTGCTGCTGTCATGCTGTGTTTGCCTTTAAAGCCAGGA	660
DB	601	TCCTGACAATCGATAGGTACTGCTGCTGTCATGCTGTGTTTGCCTTTAAAGCCAGGA	660
QY	661	CGGTCACCTTTGGGTTGGTGAAAGTGTGATCACTTGGGTGGTGGCTGTGTTGGGTCTC	720
DB	661	CGGTCACCTTTGGGTTGGTGAAAGTGTGATCACTTGGGTGGTGGCTGTGTTGGGTCTC	720
QY	721	TCCAGAGATCATCTTTACAGATCTCAAAAGAGGCTGTTCATTACACTGCACTGCTC	780
DB	721	TCCAGAGATCATCTTTACAGATCTCAAAAGAGGCTGTTCATTACACTGCACTGCTC	780
QY	781	ATTTTTCCAT-----ACATTAAAGATGATCACT	808
DB	781	ATTTTTCCATCACTCAGTATCAATCTGGAAGATTTCCAGACATTTAAAGATAGTCACT	840
QY	809	TGGGCTGGTCTGGCGCTGTTGTTCATGGTCACTCTGCTACTCGGGAATCCTTAAAACTC	868
DB	841	TGGGCTGGTCTGGCGCTGTTGTTCATGGTCACTCTACTCGGGAATCCTTAAAACTC	900
QY	869	TGCTTCGGTCTCCAAATGAGAGAGAGGCACAGGCTGTGAGGCTTATCTTCACCATCA	928
DB	901	TGCTTCGGTCTCCAAATGAGAGAGAGGCACAGGCTGTGAGGCTTATCTTCACCATCA	960
QY	929	TGATTGTATTATTTCTCTCTGGGCTCCCTCAACAATTTGCCCTCTCTCTGAAACCTGCC	988
DB	961	TGATTGTATTATTTCTCTCTGGGCTCCCTCAACAATTTGCCCTCTCTCTGAAACCTGCC	1020
QY	989	AGGAATCTTTTGGCTTGAATAATTTGCAGTAGTCTTAAACAGTTGGACAAGCTATGCAGG	1048
DB	1021	AGGAATCTTTTGGCTTGAATAATTTGCAGTAGTCTTAAACAGTTGGACAAGCTATGCAGG	1080
QY	1049	TGACAGAGACTCTTGGGATGACGCACTGCTGCATCAACCCCATCATCTATGTCCTTTGTCG	1108
DB	1081	TGACAGAGACTCTTGGGATGACGCACTGCTGCATCAACCCCATCATCTATGTCCTTTGTCG	1140
QY	1109	GGGAGAGTTTCAAAACTACCTCTTATGTTCTTTTCCAAAGACACATTCGCAAAACGCTCT	1168

RESULT 10

US-09-796-202-2

Sequence 2, Application US/09796202

; Patent No. US20020068813A1

; GENERAL INFORMATION:

; APPLICANT: Dragic, Tatjana

APPLICANT: OLSON, WILLIAM

: AFFILIANT: OISON, WILLIAM

: TITLE OF INVENTION: SULFATED CCR5 PEPTIDES FOR HIV-1 INFECTION

FILE OF INVENTION: SULFATED CCRS  
FILE REFERENCE: 2048/61010/TPW/SHS

FILE REFERENCE: 2048/61010/JPW/SHS  
CURRENT APPLICATION NUMBER: US/09/796.202

; CURRENT APPLICATION NUMBER: US/0  
 : CURRENT FILING DATE: 2001-02-28

; CURRENT FILING DATE: 2001-02-  
 ; NUMBER OF SEQ ID NOS: 17

```

; NUMBER OF SEQ ID NOS: 1/
: SOFTWARE: PatentIn version 3.0

```

```

; SOFTWARE: PatentIn version 3.0
- SEQ ID NO 2

```

```

; SEQ ID NO 2
: LENGTH: 1376

```

```
; LENGTH: 13
; TYPE: DNA
```

TYPE: DNA  
ORGANISM: human

; ORGANISM: h  
NC-09-706-303-3

Query Match 90.3%; Score 1302; DB 10; Length 1376;

Best Local Similarity 97.78; Pred. No. 0;

Best local similarity 97.7%, FPR: NO: 0,  
Matches 1344: Conservative 0: Mismatches 32:  
Indels 32:

QY	1	GAATTC	CCCCCA	CACAGAC	CAAGCTCT	CCATCT	AGTGGAC	AGGGAAG	CTAGCAG	CAAAACCT	60																			
Db	1	GAATTC	CCCCCA	CACAGAC	CAAGCTCT	CCATCT	AGTGGAC	AGGGAAG	CTAGCAG	CAAAACCT	60																			
QY	61	TCCCTT	CAC	TAAAC	CTTCAT	CTGCTT	GGCC	AAAGAG	AGTTAA	TTCATG	TAGACATC	120																		
Db	61	TCCCTT	CAC	TAAAC	CTTCAT	CTGCTT	GGCC	AAAGAG	AGTTAA	TTCATG	TAGACATC	120																		
QY	121	TATGTAG	GCAAT	TAAAC	AGCTATT	GATG	TATA	AAACAG	TITG	CATT	CATGGAGG	CAACT	180																	
Db	121	TATGTAG	GCAAT	TAAAC	AGCTATT	GATG	TATA	AAACAG	TITG	CATT	CATGGAGG	CAACT	180																	
QY	181	AAATCA	CTT	TAG	GACTTT	TATA	AAAGAT	CACTTT	TAT	TATATG	CACAGG	TGGAA	CAAGA	240																
Db	181	AAATCA	CTT	TAG	GACTTT	TATA	AAAGAT	CACTTT	TAT	TATATG	CACAGG	TGGAA	CAAGA	240																
QY	241	TGGATT	AT	CAAG	TG	CAAGT	CC	CAATCT	AT	GACAT	CAAT	TAT	TATAC	TG	GAG	CCCT	TGCC	300												
Db	241	TGGATT	AT	CAAG	TG	CAAGT	CC	CAATCT	AT	GACAT	CAAT	TAT	TATAC	TG	GAG	CCCT	TGCC	300												
QY	301	AAAAAT	CAAT	TGT	GAAG	CAAA	TGG	CAG	CGCG	CTCT	CGCT	CCG	CT	TACT	CA	CTG	TGTGT	360												
Db	301	AAAAAT	CAAT	TGT	GAAG	CAAA	TGG	CAG	CGCG	CTCT	CGCT	CCG	CT	TACT	CA	CTG	TGTGT	360												
QY	361	TCATCT	TTG	TTTTG	TGG	CAAC	ATG	CT	TGGT	CAT	CTC	CT	CA	TC	CT	GATA	AAAT	TG	CAAA	AGGC	420									
Db	361	TCATCT	TTG	TTTTG	TGG	CAAC	ATG	CT	TGGT	CAT	CTC	CT	CA	TC	CT	GATA	AAAT	TG	CAAA	AGGC	420									
QY	421	TGAAG	CA	CTG	ACT	GAT	CT	TAC	CT	G	GC	CA	CT	CT	G	GC	CA	CT	CT	G	TTT	TC	TT	C	480					
Db	421	TGAAG	CA	CTG	ACT	GAT	CT	TAC	CT	G	GC	CA	CT	CT	G	GC	CA	CT	CT	G	TTT	TC	TT	C	480					
QY	481	TTACT	GT	CC	CT	TC	T	GG	CT	CT	CA	T	GT	CT	G	CG	CC	CA	CT	G	GG	CA	CT	T	T	GC	CA	AT	GT	540







Db 1041 GGAATCTTTGGCTGAATAATTTGAGTAGCTTAACAGCTTGGACCAAGCTATGCGAGT 1100  
QY 1050 GACAGAGACTCTTGGATGAGCAGCTGTGATCAACCCCATCATCTATGCTTTGTGG 1109  
Db 1101 GACAGAGACTCTTGGATGAGCAGCTGTGATCAACCCCATCATCTATGCTTTGTGG 1160  
QY 1110 GGAGAGTTTCAGAACTACCTCTTCTTCTTCCAAAGACACATTTGCCAAAGCTTTG 1169  
Db 1161 GGAGAGTTTCAGAACTACCTCTTCTTCTTCCAAAGACACATTTGCCAAAGCTTTG 1220  
QY 1170 CAATGCTCTTCTTCTTCTTCCAAAGACACATTTGCCAAAGCTTTG 1229  
Db 1221 CAATGCTCTTCTTCTTCTTCCAAAGACACATTTGCCAAAGCTTTG 1280  
QY 1230 ATCCACTGGGAGCAGGAATATCTGTGGCTTGTGACAGGACTCAAGTGGGCTGGTGA 1289  
Db 1291 ATCCACTGGGAGCAGGAATATCTGTGGCTTGTGACAGGACTCAAGTGGGCTGGTGA 1340  
QY 1290 CCCAGTCAGAGTTGTGCACATGCTTGTGATTTTTCATACAGAGCTGGGCTGGGCTGGTGA 1349  
Db 1341 CCCAGTCAGAGTTGTGCACATGCTTGTGATTTTTCATACAGAGCTGGGCTGGGCTGGTGA 1400  
QY 1350 GGAGGCTCTTTT 1362  
Db 1401 GGAAGAGTCTTT 1413

RESULT 13  
US-09-195-662A-1  
; Sequence 1, Application US/09195662A  
; Patent No. US20020076745A1  
; GENERAL INFORMATION:  
; APPLICANT: Li, Yi  
; APPLICANT: Ruben, Steven, M.  
; TITLE OF INVENTION: Human G-Protein Chemokine Receptor HDGNR10 (CCR5 Receptor)  
; FILE REFERENCE: 1488.1150002  
; CURRENT APPLICATION NUMBER: US/09/195,662A  
; CURRENT FILING DATE: 1998-11-18  
; PRIOR APPLICATION NUMBER: 08/466,343  
; PRIOR FILING DATE: 1995-06-06  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 1414  
; TYPE: DNA  
; ORGANISM: Artificial Sequence: Genomic  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (259)..(1314)  
; OTHER INFORMATION: Description of Artificial Sequence: Genomic  
US-09-195-662A-1

Query Match 89.7%; Score 1293.4; DB 10; Length 1414;  
Best Local Similarity 96.6%; Pred. No. 0;  
Matches 1345; Conservative 0; Mismatches 16; Indels 32; Gaps 1;

QY 2 AATTCGCCCAAGCAGCAAGCTCTCCATCTAGTGGACAGGAAAGTATGACGAAACCTT 61  
Db 21 AATTCGCCCAAGCAGCAAGCTCTCCATCTAGTGGACAGGAAAGTATGACGAAACCTT 80  
QY 62 CCCTTCACATAAAGCTCTATGCTTGGCCAAAGAGAGATTAATTCATATGACAGATCT 121  
Db 81 CCCTTCACATAAAGCTCTATGCTTGGCCAAAGAGAGATTAATTCATATGACAGATCT 140  
QY 122 ATGAGGCAATTAAGCACTTATGATTAATTAAGCACTTATGATTAATTAAGCACTTAT 181  
Db 141 ATGAGGCAATTAAGCACTTATGATTAATTAAGCACTTATGATTAATTAAGCACTTAT 200  
QY 182 AATACATCTAGGACTTTATAAAGATCACATTTTATTTATGACAGGCTGGAACAGAT 241  
Db 201 AATACATCTAGGACTTTATAAAGATCACATTTTATTTATGACAGGCTGGAACAGAT 260  
QY 242 GGAATTACAGTGTCAAGTCCCAATCTATGACATCAATTAATTAATACATCGGAGCCCTGCCA 301

Db 261 GGAATTACAGTGTCAAGTCCCAATCTATGACATCAATTAATTAATACATCGGAGCCCTGCCA 320  
QY 302 AAAAATCAATGTGAAGCAAAATCGAGCCCGCTCTCGCTCCGCTCTACTCTACTCTGCTGTT 361  
Db 321 AAAAATCAATGTGAAGCAAAATCGAGCCCGCTCTCGCTCCGCTCTACTCTACTCTGCTGTT 380  
QY 362 CATCTTTGGTTTGGGGGACACATGCTGGTCACTCTCATCTCTATCTCTGATTAATGCAAAAGCT 421  
Db 381 CATCTTTGGTTTGGGGGACACATGCTGGTCACTCTCATCTCTGATTAATGCAAAAGCT 440  
QY 422 GAAAGACATGACATGACATCTACCTGCTCAACTGGCCATCTCTGACCTGCTGTTTCTCTCT 481  
Db 441 GAAAGACATGACATGACATCTACCTGCTCAACTGGCCATCTCTGACCTGCTGTTTCTCTCT 500  
QY 482 TACTGTCCCTCTCTGGGCTCACTATGCTGCCGCCAGCTGGGACTTTGGAATFACAATGTG 541  
Db 501 TACTGTCCCTCTCTGGGCTCACTATGCTGCCGCCAGCTGGGACTTTGGAATFACAATGTG 560  
QY 542 TCAACTCTTACAGGCTCTATTTTATAGGCTTCTCTCTGGAATCTTCTTCAATCATCTCT 601  
Db 561 TCAACTCTTACAGGCTCTATTTTATAGGCTTCTCTCTGGAATCTTCTTCAATCATCTCT 620  
QY 602 CCTGACAAATCGATAGGTACCTGGCTGCTGCCATGCTGTTTGGCTTTAAAGCCAGGAC 661  
Db 621 CCTGACAAATCGATAGGTACCTGGCTGCTGCCATGCTGCTGTTTGGCTTTAAAGCCAGGAC 680  
QY 662 GGTCACTTTGGGCTGGTGAACATGTAATCTGCTGGTGGTGGCTGCTGTTTGGCTCTCT 721  
Db 681 GGTCACTTTGGGCTGGTGAACATGTAATCTGCTGGTGGTGGCTGCTGTTTGGCTCTCT 740  
QY 722 CCCAGGAATCATCTTTACAGATCTCAAAAGAGGCTCTCAATACACCTGACGCTCTCA 781  
Db 741 CCCAGGAATCATCTTTACAGATCTCAAAAGAGGCTCTCAATACACCTGACGCTCTCA 800  
QY 782 TTTTCCAT-----ACATTTAAAGATAGTCACTTT 809  
Db 801 TTTTCCATACAGTCAGTATCAATCTGGAAGAAATTTCCAGACATTAAGATAGTCACTTT 860  
QY 810 GGGCTGGTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 869  
Db 861 GGGCTGGTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 920  
QY 870 GCTTCGGTGTGCAATGAGAAAGAGGACAGGCTGTGAGGCTTATCTTCAACATCAT 929  
Db 921 GCTTCGGTGTGCAATGAGAAAGAGGACAGGCTGTGAGGCTTATCTTCAACATCAT 980  
QY 930 GATTTGTTTATTTCTCTTCTGGGCTCCCTTACACATTTGCTCTTCTCTGAAACACCTTCCA 989  
Db 981 GATTTGTTTATTTCTCTTCTGGGCTCCCTTACACATTTGCTCTTCTCTGAAACACCTTCCA 1040  
QY 990 GGAATTTCTTTGGCTGCAATATTTGACATGCTTACAGGCTTGGACCAAGCTATGCAAGT 1049  
Db 1041 GGAATTTCTTTGGCTGCAATATTTGACATGCTTACAGGCTTGGACCAAGCTATGCAAGT 1100  
QY 1050 GACAGACACTCTTGGGATGAGCAGCTGCTGCATCAACCCCATCATCTATGCTGCTTTGCGG 1109  
Db 1101 GACAGACACTCTTGGGATGAGCAGCTGCTGCATCAACCCCATCATCTATGCTGCTTTGCGG 1160  
QY 1110 GGAGAGTTTCAGAACTACCTCTTCTTCTTCCAAAGACACATTTGCCAAAGCTTTG 1169  
Db 1161 GGAGAGTTTCAGAACTACCTCTTCTTCTTCCAAAGACACATTTGCCAAAGCTTTG 1220  
QY 1170 CAATGCTCTTCTTCTTCTTCCAGCAAGGCTCCGAGGAGCAAGCTGATGTTACACCG 1229  
Db 1221 CAATGCTCTTCTTCTTCTTCCAGCAAGGCTCCGAGGAGCAAGCTGATGTTACACCG 1280  
QY 1230 ATCCACTGGGAGCAGGAATATCTGTGGCTTGTGACAGGACTCAAGTGGGCTGGTGA 1289  
Db 1281 ATCCACTGGGAGCAGGAATATCTGTGGCTTGTGACAGGACTCAAGTGGGCTGGTGA 1340  
QY 1290 CCCAGTCAGAGTTGTGCACATGCTTGTGATTTTTCATACAGGCTGGGCTGGGCTGGTGA 1349



; CURRENT APPLICATION NUMBER: US/09/502,783A  
 ; CURRENT FILING DATE: 2001-08-23  
 ; PRIOR APPLICATION NUMBER: 08/466,343  
 ; PRIOR FILING DATE: 1995-06-06  
 ; NUMBER OF SEQ ID NOS: 9  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 1  
 ; LENGTH: 1414  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (259)..(1314)  
 ; US-09-502-783A-1

Query Match 89.7%; Score 1293.4; DB 10; Length 1414;  
 Best Local Similarity 96.6%; Pred. No. 0;  
 Matches 1345; Conservative 0; Mismatches 16; Indels 32; Gaps 1;

QY	2	AATTCCTCCCAAGAGCAAGCTCTCCATCTAGTGGACAGGAGCAAGCTAGCAGCAACCTT	61
DB	21	ATTCCTCCCAAGAGCAAGCTCTCCATCTAGTGGACAGGAGCAAGCTAGCAGCAACCTT	80
QY	62	CCCTTCACCTACAAACTTCACTGCTTGGCCAAAGAGAGAGTTAATTCATGTAGACATCT	121
DB	81	CCCTTCACCTACAAACTTCACTGCTTGGCCAAAGAGAGAGTTAATTCATGTAGACATCT	140
QY	122	ATGTAGGCAATTAACCACTATGATGATTAACAGTTGTCATTCATGAGGAGCAACTA	181
DB	141	ATGTAGGCAATTAACCACTATGATGATTAACAGTTGTCATTCATGAGGAGCAACTA	200
QY	182	AATACATCTAGGACTTTAATAAGATCACCTTTTATATGCACAGGGTGGACAAGAT	241
DB	201	AATACATCTAGGACTTTAATAAGATCACCTTTTATATGCACAGGGTGGACAAGAT	260
QY	242	GGATTAATCAAGTGTCAAGTCCCAATCATGACATCAATTAATATATATATATATATAT	301
DB	261	GGATTAATCAAGTGTCAAGTCCCAATCATGACATCAATTAATATATATATATATATAT	320
QY	302	AAAAATCAATGTGAAGCAAAATGCGAGCCGCTCTGCTCCGCTCTACTACTCTGCTGT	361
DB	321	AAAAATCAATGTGAAGCAAAATGCGAGCCGCTCTGCTCCGCTCTACTACTCTGCTGT	380
QY	362	CATCTTTGGTTTTGTGGCAACATGCTGCTCATCTCATCTGATAAAGTCAAAAGGCT	421
DB	381	CATCTTTGGTTTTGTGGCAACATGCTGCTCATCTCATCTGATAAAGTCAAAAGGCT	440
QY	422	GAAGACATGACTGACATCTACCTGCTCAACCTGGCCATCTGACCTGTTTTCTCTCT	481
DB	441	GAAGACATGACTGACATCTACCTGCTCAACCTGGCCATCTGACCTGTTTTCTCTCT	500
QY	482	TACTGTCCCTTCTGGGCTCAGTATGCTGCGCCGACGTGGACTTTGGAATACAAATG	541
DB	501	TACTGTCCCTTCTGGGCTCAGTATGCTGCGCCGACGTGGACTTTGGAATACAAATG	560
QY	542	TCACATCTTTGAGGGCTCPATTTATAGGCTTCTCTGGAATCTTCTTCATCATCT	601
DB	561	TCACATCTTTGAGGGCTCPATTTATAGGCTTCTCTGGAATCTTCTTCATCATCT	620
QY	602	CCCTGCAATTCAGTAGTACCTGGCTGCTGCTCCATGCTGCTTTTAAAGCCAGGAC	661
DB	621	CCCTGCAATTCAGTAGTACCTGGCTGCTGCTCCATGCTGCTTTTAAAGCCAGGAC	680
QY	662	GGTCACCTTTGGGTTGGTGAAGTGTGATCAGTTGGTGGTGGTGGTGGTGGTGGTGGT	721
DB	681	GGTCACCTTTGGGTTGGTGAAGTGTGATCAGTTGGTGGTGGTGGTGGTGGTGGTGGT	740
QY	722	CCCAGGAATCATCTTTACAGATCTCAAAAGAGAGTCTTCATTCATTCAGCTCTCA	781
DB	741	CCCAGGAATCATCTTTACAGATCTCAAAAGAGAGTCTTCATTCATTCAGCTCTCA	800
QY	782	TTTTCCAT-----ACATTAAGATAGTCACTT	809

DB	801	TTTTCCATCAGTCAGTATCAATTTCTGGAGAAATTTCCAGACATTTAAAGATGATCACTT	860
QY	810	GGGGCTGGTCTCGCGCTGCTGTGTCATGTCATGTCATGTCATGTCATGTCATGTCAT	869
DB	861	GGGGCTGGTCTCGCGCTGCTGTGTCATGTCATGTCATGTCATGTCATGTCATGTCAT	920
QY	870	CTTTCCGCTGTCGAATGAGAAAGAGAGGACAGGCTGTGAGGCTTATCTTCACCATCAT	929
DB	921	CTTTCCGCTGTCGAATGAGAAAGAGAGGACAGGCTGTGAGGCTTATCTTCACCATCAT	980
QY	930	GATTTGTTTATTTCTCTCTGGGCTCCCTACAAACATTTCTCTCTCTCTCTCTCTCTCT	989
DB	981	GATTTGTTTATTTCTCTCTGGGCTCCCTACAAACATTTCTCTCTCTCTCTCTCTCTCT	1040
QY	990	GGAATTTTGGCGCTGAATAATTCAGTACTCTTAACAGTTTGGACCAAGCTATGCAAGT	1049
DB	1041	GGAATTTTGGCGCTGAATAATTCAGTACTCTTAACAGTTTGGACCAAGCTATGCAAGT	1100
QY	1050	GACAGACACTTGGGATGACGCACTGCTGCATCAACCCCATCATCTATGCCCTTGTGCG	1109
DB	1101	GACAGACACTTGGGATGACGCACTGCTGCATCAACCCCATCATCTATGCCCTTGTGCG	1160
QY	1110	GGAGAAGTTTCAGAACTACCTTTAGTCTTTCTTCCAAAGCACATTCGCAACGCTCTG	1169
DB	1161	GGAGAAGTTTCAGAACTACCTTTAGTCTTTCTTCCAAAGCACATTCGCAACGCTCTG	1220
QY	1170	CAATGCTGTTCTATTTTCCAGGAGAGGCTCCCGAGGAGCAAGCTCAGTTTACACCG	1229
DB	1221	CAATGCTGTTCTATTTTCCAGGAGAGGCTCCCGAGGAGCAAGCTCAGTTTACACCG	1280
QY	1230	ATCCACTGGGGAGCAGGAAATATCTGTGGGCTTGTGACAGGACTCAAGTGGGCTGTGA	1289
DB	1281	ATCCACTGGGGAGCAGGAAATATCTGTGGGCTTGTGACAGGACTCAAGTGGGCTGTGA	1340
QY	1290	CCCAGTCAGAGTTGTGCACATGGCTTAGTTTTCATACAGCCCTGGGCTGGGCTGGT	1349
DB	1341	CCCAGTCAGAGTTGTGCACATGGCTTAGTTTTCATACAGCCCTGGGCTGGGCTGGT	1400
QY	1350	GGAGTCTTTTTT	1362
DB	1401	GGAGAGGCTTTT	1413

Search completed: June 8, 2003, 10:50:57  
 Job time : 228 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 06:22:47 ; Search time 2237 Seconds  
(without alignments)  
10439.827 Million cell updates/sec

Title: US-09-938-703-3

Perfect score: 1442

Sequence: 1 GAATTCCTCCCAACAGAGCA.....AGTAGATTAGATCCGATTC 1442

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estm:\*  
5: em\_estov:\*  
6: em\_estpi:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_man:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	654	45.4	807	13	BI764263
2	360.6	25.0	487	10	BE656336
3	327.8	22.7	542	9	AA547303
4	322	22.3	452	9	AI851510
5	315.6	21.9	480	10	AW657263
6	280.8	19.5	447	9	AA671573

7	272.2	18.9	907	9	AL552677
8	268	18.6	1074	14	BM917063
9	222.6	15.4	952	14	BM917763
10	222.4	15.4	789	12	BG205056
11	211.6	14.7	884	12	BG182330
12	211.6	14.7	1167	14	BQ053936
13	207.4	14.4	876	13	BI906283
14	204.2	14.0	672	10	BB638766
15	202.4	14.0	745	12	BG204024
16	190.4	13.2	869	12	BF119225
17	187.4	13.0	568	12	BF193051
18	174	12.1	533	12	BF193021
19	170.2	11.8	669	13	BI393893
20	167.2	11.8	934	12	BG460984
21	166.4	11.5	639	10	BB629533
22	164.6	11.4	469	17	BM4215
23	162.4	11.3	863	13	BI661279
24	157.8	10.9	646	10	BB615654
25	157.8	10.9	662	10	BB644135
26	157.8	10.9	665	10	BB628567
27	157.8	10.9	723	14	BM951933
28	157.8	10.9	787	9	AU080004
29	157.8	10.9	936	12	BF119806
30	157.8	10.9	965	14	BQ944555
31	157.6	10.9	551	12	BF081124
32	156.2	10.8	600	13	BI987229
33	154.4	10.7	451	10	AM658666
34	149.6	10.4	3005	11	AK019478
35	146	10.1	669	10	BB642553
36	144.4	10.0	867	13	BI106339
37	139.4	9.7	650	10	BB630138
38	138.6	9.6	2111	11	AK007808
39	138.4	9.6	450	14	W04836
40	136.6	9.5	780	13	BI908127
41	136.4	9.5	770	9	AJ452673
42	135	9.4	558	12	BE750883
43	135	9.4	659	17	AZ667240
44	130	9.0	1098	17	CNS0586L
45	129.8	9.0	417	17	AQ341822

#### ALIGNMENTS

RESULT 1  
BI764263  
LOCUS  
DEFINITION 603045953f1 NIH\_MGC\_116 Homo sapiens cDNA clone IMAGE:5186388 5',  
807 bp mRNA linear EST 25-SEP-2001  
mRNA sequence.  
ACCESSION BI764263  
VERSION BI764263.1 GI:15755841  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 807)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapsb-r@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM11465 row: i column: 13  
High quality sequence stop: 805.  
Location/Qualifiers  
1. .807

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE:5186388"
/clone_lib="NIH_MGC_116"
/lab_host="DH10B"
/notes="Organ: pooled colon, kidney, stomach; Vector:
PCMV-SPT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 45 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."
BASE COUNT      186 a   209 c   171 g   241 t
ORIGIN
Query Match      45.4%; Score 654; DB 13; Length 807;
Best Local Similarity 94.8%; Pred. No. 7.5e-185;
Matches 743; Conservative 0; Mismatches 5; Indels 36; Gaps 5;
QY 224 CACAGGTGGAACAGATGATTATCAAGTCCAACTATGATGATCAATATTATTA 283
D 28 CCGCGGTGGAACAGATGATTATCAAGTGGAGTCCATCTATGATCAATATTATTA 87
QY 284 TATATCGAGGCGCTGCCAAAATCAATGTGAAGCAATCGACGCGCGCTTCGTGCTC 343
D 88 TATATCGAGGCGCTGCCAAAATCAATGTGAAGCAATCGACGCGCGCTTCGTGCTC 147
QY 344 GCTCTACTCATGCTGTTCACTCTTGTGTTGGGCAACATGCTGGTCATCTCATCTT 403
D 148 GCTCTACTCATGCTGTTCACTCTTGTGTTGGGCAACATGCTGGTCATCTCATCTT 207
QY 404 GATAACTGCAAAAGGCTGAAGAGCATGACTGACATCTACTGCTCAACCTGGCCATCTC 463
D 208 GATAACTGCAAAAGGCTGAAGAGCATGACTGACATCTACTGCTCAACCTGGCCATCTC 267
QY 464 TGACCTGTTTTCCTTCTACTGTCCTTCTGGGCTCAGTATGTCGGCGGCGAGTGGGA 523
D 268 TGACCTGTTTTCCTTCTACTGTCCTTCTGGGCTCAGTATGTCGGCGGCGAGTGGGA 327
QY 524 CTTTGGAAATACAAATGCTCAACTCTTGACAGGCTCTATTATATAGGCTTCTTCTG 583
D 328 C-TTGGAAATACAAATGCTCAACTCTTGACAGGCTCTATTATATAGGCTTCTTCTG 386
QY 584 AATCTTCTCATCATCTCTCTGCAATCGATAGTACCTGGCTGCTGCTGCTGCTGTT 643
D 387 AATCTTCTCATCATCTCTCTGCAATCGATAGTACCTGGCTGCTGCTGCTGCTGTT 446
QY 644 TCGTTTAAAGCCAGGAGGTCACCTTTGGGCTGTGACAGTGATCAGTGGGTGGT 703
D 447 TCGTTTAAAGCCAGGAGGTCACCTTTGGGCTGTGACAGTGATCAGTGGGTGGT 506
QY 704 GCGTGTGTTGGCTCTCCGAGGAATCATCTTTACAGATCTCAAAAAGAGTCTTCA 763
D 507 GCGTGTGTTGGCTCTCCGAGGAATCATCTTTACAGATCTCAAAAAGAGTCTTCA 566
QY 764 TTACACCTGACGCTCATTTTCAT-----AC 791
D 567 TTACACCTGACGCTCATTTTCATTTCCATACAGTCAGTATCAATTTCTGGAAGATTTCCAGAC 626
QY 792 ATTTAAGATAGTACTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 851
D 627 ATTTAAGATAGTACTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 686
QY 852 GGGATCCTTAAAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 911
D 687 GGGATCCTTAAAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 745
QY 912 GCTTATCTTCCATCATCATGATGTTTATTTCTCTTCTGGGCTCCCTACAAATATGTCCT 971
|||||
```

```
Db 746 GCTTATCTTCCATCATCATGATGTTTATTTCTCTTCTGGGCT-CCTACACATGTCCT 803
QY 972 TCTC 975
D 804 TCTC 807
|||||
RESULT 2
BE656336
LOCUS      BE656336
DEFINITION UI-M-BH0-aju-h-06-0-UI.r1 NIH_BMAP_M.S1 Mus musculus cDNA clone
            487 bp mRNA linear EST 06-SEP-2000
ACCESSION BE656336
VERSION   BE656336.1 GI:9982249
KEYWORDS  EST.
SOURCE    house mouse.
ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 487)
AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
            discovery
JOURNAL    Genome Res. 6 (9), 791-806 (1996)
MEDLINE    9704477
COMMENT    Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
            20892-9643, USA
            Tel: 301 443 1706
            Fax: 301 443 9890
            Email: mst@mail.nih.gov
            cDNA Library Preparation: M.B. Soares Lab Clone distribution:
            Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
            should be noted that Bento Soares is generating a small number of
            additional specialized non-redundant arrays of BMAP cDNAs whose
            availability will be considered under appropriate and limited
            collaborative arrangements
            Seq primer: M13 Reverse.
            Location/Qualifiers
                1..487
                /organism="Mus musculus"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone="UI-M-BH0-aju-h-06-0-UI"
                /clone_lib="NIH_BMAP_M.S1"
                /dev_stage="27-32 days"
                /lab_host="DHL0B (Life Technologies)"
                /notes="Vector: pMT3D-Pac (Pharmacia) with a modified
                polylinker; Site_1: Not I; Site_2: Eco RI; The
                NIH_BMAP_M.S1 library is a subtracted library derived from
                a mixture of normalized libraries from ten regions of the
                mouse brain (cerebellum, brain stems, olfactory bulbs,
                hypothalamus, cortex, amygdala, basal ganglia, pineal
                gland, striatum, hippocampus). The driver used for
                subtraction consisted of a pool of 20,000 cDNA clones
                obtained from non-normalized and normalized libraries of
                these ten regions of the mouse brain."
BASE COUNT      115 a   115 c   103 g   154 t
ORIGIN
Query Match      25.0%; Score 360.6; DB 10; Length 487;
Best Local Similarity 84.6%; Pred. No. 8.3e-97;
Matches 405; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 264 ATCTATGATCAATATTATTATATACATCGGAGCCCTGCCAAAAATCAATGTGAAGCAATC 323
D 9 AGCTATGATCATGATATGATGTCAGCACCTGCCAAAAATCAATGTGAAGCAAT 68
QY 324 CGAGCCCGCTCTGCTCGCTCTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 383
D 69 GGGGCTAGCTCTCTGCCCCCTACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 128
|||||
```



Db 91 CAGCAGAACAAATCCTGATCGTGCAGCTATACCCGATATCCAGGAGAACATGAA 32  
 QY 1249 ATATCTGTGGGCTT 1262  
 Db 31 GTTCTACTGGTTT 18  
 RESULT 5  
 LOCUS AW657263 480 bp mRNA linear EST 25-APR-2001  
 DEFINITION 109996 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.  
 ACCESSION AW657263  
 VERSION AW657263.1 GI:7423089  
 KEYWORDS EST.  
 SOURCE COW.  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovidae; Bovinae; Bos.  
 REFERENCE 1 (bases 1 to 480)  
 AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,  
 Casas,E., Wray,J.B., White,J., Cho,J., Fahrénkrug,S.C., Bennett  
 ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,  
 Perle,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and  
 Keefe,J.W.  
 TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA  
 libraries and construction of a gene index for cattle  
 JOURNAL Genome Res. 11 (4), 626-630 (2001)  
 MEDLINE 21180013  
 COMMENT Contact: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390  
 Email: smith@email.marc.usda.gov  
 Single pass sequencing. Bases called and alt\_trimmed with phred  
 v0.980304.e. Vector identified by cross\_match with the -mnscore 18  
 and -minmatch 12 options.  
 PCR Primers  
 FORWARD: AGGAACAGCTATGACCAT  
 BACKWARD: GTTTCCTCAGTCACGACG  
 Plate: 100 row: 0 column: 13  
 Seq primer: ATTAGTGACACTATAG.  
 FEATURES  
 source  
 location/Qualifiers  
 1. 480  
 /organism="Bos taurus"  
 /db\_xref="taxon:9913"  
 /clone\_lib="MARC 1BOV"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV SPOR6; Site\_1: NotI; Site\_2: SalI;  
 Library made from pooled tissue from lymph node, ovary,  
 fat, hypothalamus, and pituitary."  
 BASE COUNT 110 a 130 c 112 g 128 t  
 ORIGIN  
 Query Match 21.9%; Score 315.6; DB 10; Length 480;  
 Best Local Similarity 82.9%; Pred. No. 2.7e-83;  
 Matches 398; Conservative 0; Mismatches 49; Indels 33; Gaps 2;  
 QY 663 GTCACTTTGGGTGGTGACAAAGTGATCACTTGGGTGGGTGTTGGTCTCTC 722  
 Db 1 GTCACTTTGGGTGGGTGGTGACAAAGTGATCACTTGGGTGGGTGTTGGTCTCTC 60  
 QY 723 CCAGGAATCATCTTTACCAAGATCTCAAAAAGAGGTCTTATACATCGTCAGCTCTCAT 782  
 Db 61 CCAGGAATCATCTTTACCAAAATCCAAAAGAGAGGCTCTCGTCATCATCAGCCACAT 120  
 QY 783 TTTCATACAC-----TTAAGATGATCATCTTG 810  
 Db 121 TTCCCATCCAGTCAGTATCATCTTCGTGAAGAAATTCCAAACTTTAAGATGATCATCTTG 180  
 QY 811 GGGCTGGTCTCGCGCTGTTGTCTGATCGTCATCTGCTACTCGCGGAATCCTTAAACATCTTG 870





vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com

BASE COUNT 189 a 238 c 212 g 267 t 1 others

ORIGIN

Query Match 18.9%; Score 272.2; DB 9; Length 907;  
Best Local Similarity 61.6%; Pred. No. 4e-70;  
Matches 507; Conservative 0; Mismatches 278; Indels 38; Gaps 3;

QY 388 TGGTCATCTCTCTGATTAACGCAAGGCTGAAGAGCATGACATCTACTGTC 447  
Db 1 TGGTGGCTGGTCTGTGCAATACAGAGGCTAAACATGACCATCTACTGTC 60  
QY 448 TCAACCTGGCCTCTCTGACCTGTTTCTTCTTCTCTCTCTCTCTCTCTCT 504  
Db 61 TGAACCTGGCCTCTCTGACCTGTTTCTTCTTCTCTCTCTCTCTCTCTCT 120  
QY 505 ATGCTGCGCCGCGGAGTGGAGCTTGGAAATACAAATGTCACACTCTTTCAGAGGCTCTATT 564  
Db 121 AGTTGAAGAGTACTGGGTTTTTGGTGATGCCATGTAAGATCTCTCTCTCTCTCTCTCT 180  
QY 565 TTATAGGCTTCTCTCTGAAATCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 624  
Db 181 ACACAGCTGTGACAGCAGATCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240  
QY 625 CTGTGCT 684  
Db 241 CCATCGTCCACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
QY 685 GTGTGATCATCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 744  
Db 301 GCATCATCATCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360  
QY 745 CTCACAAAGAGGCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 786  
Db 361 CCCATGGATCTACTCACCACACCTGACGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420  
QY 787 -----CATCATTAAGATAGTCATCTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 832  
Db 421 AGTGAAGCTGTTTTCAGGCTCTGAACTGAACCTCTTGGGCTGGTATTGGCTTTGTTGG 480  
QY 833 TCATGCTCATCTGCTACTCGGATCTTAAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 892  
Db 481 TCATGATCATCTGCTACACGGGATTAAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
QY 893 AGAGGCACAGGCTGTGAGGCTTAATCTCACCATCATGATGTTTATTTCTTCTTCTG 952  
Db 541 A---ATCCAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 597  
QY 953 CTGCTACACATCTGCT 1012  
Db 598 CCCCCTACAAATTTGCTATCTATCTTCTGTTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 657  
QY 1013 GCATAGCTCTACAGGTTGGACCAAGCTATGAGGTGACAGAGCTCTTGGGATGACGC 1072  
Db 658 GTGAGCAGACAGACATTTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717  
QY 1073 ACTGCTGATCAACCCGATCATCTATGCTCTTGTGCGGGAGAGTTCAGAACTACCTCT 1132  
Db 718 ACTGCTGTCAACCCAGTATCTACGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 777  
QY 1133 TAGTCTTCTTCCAAAGACATTCGCAAGCTTCTGCAATG 1175  
Db 778 GGCAGTGTCTCCACAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 820

RESULT 8

BM917063

LOCUS

1074 bp mRNA linear EST 12-MAR-2002

## DEFINITION

AGENCOURT\_6702060 NIH\_MGC\_106 Homo sapiens cDNA clone IMAGE:5483554

5', mRNA sequence.

BM917063

VERSION BM917063.1 GI:19367442

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1074)

NIH-MGC http://mgs.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaaps-r@mail.nih.gov

Tissue Procurement: Dr. Daniel McVicar, DBS/MCI

cDNA Library Preparation: Rubin Laboratory

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLCM2011 row: g column: 11

High quality sequence stop: 718.

## FEATURES

Location/Qualifiers

1..1074

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:5483554"

/clone\_lib="NIH\_MGC\_106"

/tissue\_type="natural killer cells, cell line"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: blood; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

BASE COUNT 240 a 282 c 245 g 302 t 5 others

## ORIGIN

Query Match 18.6%; Score 268; DB 14; Length 1074;  
Best Local Similarity 62.08; Pred. No. 7.9e-69;  
Matches 526; Conservative 0; Mismatches 282; Indels 40; Gaps 5;

QY 294 CCTGCCAAAAATCAATGTGAAGCAAAATCGACGCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 353  
Db 93 CCGTGCCAGAGGTGAACGAGAGGCGCTTTGGGGCCCAACTGCTGCCCTCTGTACTCC 152  
QY 354 CTGGTGTTCATCTTGGTGTGGGCAACATGCTGGTCATCTCTCATCTGATTAACATGC 413  
Db 153 TTGGTATTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 212  
QY 414 AAAAGGCTGAAGAGCATGACTGACATCTTACCTGCTCAACCTGGCCATCTCTGACCTGTTT 473  
Db 213 AAGAGGCTAAAAACATGACCATCTTACCTCTCTGACCTGGCCATCTTCTGACCTGCTC 272  
QY 474 TTCTTCTTACTGTCCCTTCTGG---GCTCACTATGCTGCGGCCAGTGGAGTTCGA 530  
Db 273 TTCTCTGTACGCTTCCCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 332  
QY 531 AATCAATGTCATCTCTTACAGGCTCTATTTTATAGGCTTCTTCTCTGATCTTC 590  
Db 333 GATGCCATGTGTGAATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 392  
QY 591 TTTCATCATCTCTCTGACAAATCGATAGGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 650  
Db 393 TTTCATCATCTCTGCTGAGGATGACAGGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 452  
QY 651 AAAGCCAGAGGCTACCTTTGGGGTGGTGACAAAGTGTGATCACTTGGGTGGTGGTGGTGGTGGTGGTGGTGG 710



Tel: 216 431 9900			
Fax: 216 361 9596			
Email: scai@ethersys.com			
High quality sequence stop: 521.			
FEATURES	Location/Qualifiers		
source	1..789		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_lib="Athersys RAGE Library"		
	/cell_line="H1080"		
	/notes="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is H1080, since a random activation method was used, these sequence tags are not necessarily expressed in H1080 under normal circumstances."		
BASE COUNT	174 a 207 c 183 g 222 t		
ORIGIN			
	Query Match 15.4%; Score 222.4; DB 12; Length 789;		
	Best Local Similarity 61.5%; Pred. No. 3 4e-55;		
	Matches 426; Conservative 0; Mismatches 229; Indels 38; Gaps 3;		
QY	293 GCCTCGCCAAATAATCATGATGTAAGCAAAATPCGACGCGGCTCTCTGGCTCCGCTCTACTC 352		
Db	81 GCTCTGTGAAAAGCTGATACAGACGACCTGATGCGCCAGCTTTGTGCCCCCGCTGTACTC 140		
QY	353 ACTGGTGTTACTCTTTGGTTTGGGCAACATGCTGGTCATCTTCATCTCTGATAAACG 412		
Db	141 CTTGGTGTTACTCTGGGCTCTTGGGCAATGCTGGTGATGATGATCTCTCATAAATA 200		
QY	413 CAAAAGGCTGAAGACGATGACTGATCTACTCTGCTCAACCTGGGCCATCTCTGACCTGTT 472		
Db	201 CAGGAGGCTCGCAATATGACCAACATCTACCTGCTCAACTGCGCATTTTCGGACCTGCT 260		
QY	473 TTTCTCTTACTGTCCTCTCTGGCTGACATATGCTGCGGCC---AGTGGGACTTTGG 529		
Db	261 CTTCCTCTGCTACCTTCCATCTTGATGCACATATGTACGGGGGCATACACTGGGTTTTC 320		
QY	530 AAATCAATGTCACACTTTGACAGGGCTCTATTTTATAGGCTTCTTCTCTGGAATCT 589		
Db	321 CCATGGCATGTGTAAGCTCCTCTCAGGGTTTATCACACAGGCTTGTACACGAGATCTT 380		
QY	590 CTTTCATCTCTCTGACAACTGATAGTACTGCTGCTGTCCTCATGCTGTTGTTGCTTT 649		
Db	381 TTTTCATAATCTCTGACAACTGACAGGTACCTGGCCATTGCTCATGCTGATTTGCCCT 440		
QY	650 AAAAGCCAGGACGGTCACCTTTGGGGTGGTGCACAAAGTGTGATCATCTGGGTGTGGCTGT 709		
Db	441 TCGAGCCGGAGCTGCTACCTTTTGGTGTTCATCACCAGCATCGTCACCTGGGCGCTGGCACT 500		
QY	710 GTTTGGGCTCTCCGAGGAATCATCTTACCAGATCTCAAAAGAGAGTCTTCATTAACAC 769		
Db	501 GCTAGAGCTCTCTCTGAAATTTATCTTCTATAGAGACCGAAGTGTGTTGAAGAGACTCT 560		
QY	770 CTGCAGCTCTCAT-----TTTCCATACATATAA 797		
Db	561 NTGCAGTCTCTTTACCCAGAGGATACAGTATATAGCTGGAGGCATTTCCACACTCTGAG 620		
QY	798 GATAGTCAATCTGGGCGTGGTCTGCGCGTGTCTTGTCAAGGTCACTGCTACTCGGGAAT 857		
Db	621 AATGACCGCTCTCTGCTCTGCTCCCTCTGCTCTGCTGTTATGAGCCATCTGCTACACAGAA 680		
QY	858 CTTAAAACTCTGTTCCGGTGTGCGAAATGAGAAAGAGGACAGAGGCTGTGAGGCTTAT 917		
Db	681 CATCANACGCTGCTGAGTGCGCCAGT---AAAAAAGTACAGGCCATCCGGCTCAT 737		
QY	918 CTTCCACCATCAATGATTTTATTTCTTCTCTG 950		
Db	738 TTTTGTGTCATCATGGGCGGGTTTTCATTTCTG 770		

RESULT 11

BG182330	884 bp	mRNA	linear	EST 21-APR-2001
LOCUS	884 bp	mRNA	linear	EST 21-APR-2001
DEFINITION	RGT1196 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.			
ACCESSION	BG182330			
VERSION	BG182330.1			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1 (bases 1 to 884) Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R., Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J., Lerner, L., Costanzo, D., Hessel, J., Kozak, R., Mays, R., Smith, E., Veloso, N., Kliff, A., Hess, J., Cothren, K., Lo, K., Offenbacher, J., Danzig, J. and Ducar, M.			
TITLE	Creation of genome-wide protein expression libraries using random activation of gene expression			
JOURNAL	Nat. Biotechnol. 19 (5), 440-445 (2001)			
MEDLINE	21227151			
COMMENT	Contact: Scott J. Cain Athersys, Inc. 3201 Carnegie Ave, Cleveland, OH 44115, USA Tel: 216 431 9900 Fax: 216 361 9596 Email: scain@atersys.com High quality sequence stop: 529.			
FEATURES	Location/Qualifiers			
source	1..884			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/clone_lib="Athersys RAGE Library"			
	/cell_line="Hr1080"			
	/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is Hr1080, since a random activation method was used, these sequence tags are not necessarily expressed in Hr1080 under normal circumstances."			
BASE COUNT	191 a	231 c	206 g	256 t
ORIGIN				
Query Match	14.7%	Score 211.6;	DB 12;	Length 884;
Best local Similarity	60.9%;	Pred. No. 6.4e-52;		
Matches	436;	Conservative	0;	Mismatches 239; Indels 41; Gaps 4;
QY	293	GCCTGCGCAAAATCAATGTGAAGCAAAATCGACCGCGCTCTCGCTCGGCTCTACAC	352	
DB	82	GCTCTGTGAAGAGCTGATACCAGAGCACTGATGCCCACTTTGTGCCCGCGCTGACTC	141	
QY	353	ACTGGTGTTCATCTTTGGTTTGGGGCAACATGCTGGTGATCCTCATCTCTGATAAAGT	412	
DB	142	CTGTGTGTTCTACTGTGGGGCTCTTGGGCAATGGTGGTGATGATCTCATATAAATA	201	
QY	413	CAAAGGCTGAAGAGCACTGACTGATCTACTGCTCAACTGGCCATCTCTGACCTGTT	472	
DB	202	CAGAGGCTCCGAATATGACCAACATCTACCTGCTCAACTGGCCATTTGCGACCTGCT	261	
QY	473	TTTCTCTTCTACTGTGCCCCCTTCTGGGCTCACTACTGCTGCCGCC---	529	
DB	262	CTTCTCTGTCACCTTCCATCTGGATCCACTATGTACAGGGGGCATACTGGGTTTGG	321	
QY	530	AAATCATGTGTCACTCTTGACAGGGCTCTATTTATATAGGCTTCTTCTCTGGAACPT	589	
DB	322	CCATGGCATGTGTAAGCTCTCTCTCAGGGTTTATACACAGGCTTGATACGCGAGATCT	381	
QY	590	CTTTCATCATCTCTGACAAATCGATAGGTACTGCTGCTCATGCTGTGTTGCTT	649	
DB	382	TTTCATATCTCTGCAATCGACAGTACTGGCCATGTGCCATGCTGTTGTTGCCCT	441	
QY	650	AAAAGCCAGAGGTCACCTTTGGGGTGGTGAAGAAGTGATCACTTGGGTGGGCTGT	709	
DB	442	TCGACCCCGGACTGTCACTTTTGGTGTCACTACACAGCATGTCACCTGGGGCTGGCACT	501	





Fax: 216 361 9596

1. .745

```

/clone
/clone

```

Librarian

cell type.

express 207

14

conservative

SECRET

[Home](#)
[About Us](#)
[Services](#)
[Contact Us](#)

## DISCUSSION

|||||

GTCTACTG

GCTGAAGA

—

\_\_\_\_\_

AGTACCC

AAATGTGTC

CATGTGTA

100

—  
—  
—  
—  
—  
—

.....

התאחדות

1111

2017

**CTCTCAT**

GTGCTCTT

## TECHNICAL

057049

8.2